UNIVERSITY OF OSLO
Department of Informatics

Master Thesis

A Focal Plane Array To Model The Color Processing of the Retina

Łukasz Farian

July 21, 2014

Abstract

This thesis presents a bio-inspired color camera chip. It is the first focal plane array of photo-pixels that react to temporal color contrast of three different color spectra and thus this vision sensor emulates the color opponencies in the human retina. The array has been designed, implemented and thoroughly characterised in TSMC 90nm CMOS technology. The three different spectra are transduced into photocurrents by stacked photodiodes and temporal changes of the contrast of these three spectra are represented as asynchronous pixels event streams which are read-out off-chip by the Address Event Representation (AER) protocol. The implemented focal plane consists of 16x16 pixels. A single pixel measures 82μm x 82μm and has a fill factor of 27%. The experimental results proved that the designed sensor can discriminate between temporal color contrasts. In comparison to traditional frame-based cameras, the designed vision sensor relaxed the post processing load of a CPU preforming moving objects color recognition.

Preface

This master thesis was submitted as a long thesis contributing 60 credits to my *Master of Science* degree in *Nanoelectronics and Robotics* at the Department of Informatics, Faculty of Mathematics and Natural Sciences, University of Oslo. This work was started during the *Spring* semester 2013 and was submitted in July 2014.

I would like to express my deepest gratitude to my supervisors, Juan Antonio Leñero-Bardallo, Philipp Häfliger and Trygve Willassen for great support and the patience to answer all my questions. In moments of doubt I could always count on their help and suggestions how to solve encountered obstacles. I would also like to thank Thanh Trung Nguyen, Deyan Levski and Kristian Gjertsen Kjelgård for contribution to making the camera work. I would like to thank my sister Malwina for putting the idea of studying in Norway into my head and my parents for a boundless support.
Last but not least, I would like to express my deep thanks to Malihe Zarre Dooghabadi for making my life more colorful.

Basic Constants

Parameter	Equation	Parameter name
q	$1.602 \cdot 10^{-19} C$	electrical charge of the electron
k	$1.38 \cdot 10^{-23} JK^{-1}$	Boltzmann constant
ϵ_0	$8.854 \cdot 10^{-12} F/m$	absolute permittivity
h	$6.626 \cdot 10^{-34} Js$	Plank's constant
c	$299792458 m/s$	speed of light in vacuum
eV	$1.602176565 \cdot 10^{-19} J$	electronvolt

Table 1: Basic constants used further in this thesis.

Contents

Chapter 1

Introduction

A human body is a result of billions of years of evolution. Throughout such long time we and other natural organisms have evolved in order to adapt to the surrounding environment. The human body is a very complex system which we do not completely understand. Especially, we have still a long way to understand how our brain works. However there are fields of human body, we understand well, like a visual perception. Hence, today we try to take an inspiration from biological vision systems and create more power efficient faster and adaptive video cameras.
The point of this chapter is to give an introduction and a motivation why it is desirable to implement bio-inspired vision sensors, which are so different from the frame-based cameras used nowadays. This chapter provides a short description of the circuitry used in modern vision sensors, then it shows the difference between them and the human eye. Next, so called, *bio-inspired* vision sensors are introduced and divided into two main groups. Finally, *Goals and Outline* of this thesis are provided.

1.1 Commercial frame-based vision sensors

By the time of writing of this thesis the dominant part of modern commercial image sensors use as their sensing element an active pixel sensor (APS). Figure 1.1 shows the simplest three transistor variant of an APS. There exist improved versions of APSs but their operation principle is the same as the three transistor topology. 'Active' means here that the amplification of the small photocurrent is performed inside the pixel. Active pixels sensor consists of a photodiode, reset transistor M1, source-follower transistor M2 and a row select transistor M3.
The charge accumulates over the photodiode. The reset transistor M1 clears the accumulated charge under the photo diode by connecting it to the power supply. Next, the transistor M1 is turned off and the transistor's source (initially equal to power supply) can be now effectively discharged by the photocurrent. Transistor M2 realises a source follower amplifier which buffers the accumulated charge. Since the source follower has high input impedance, the read-out does not affect the accumulated charge. Finally, transistor M3 enables to select this pixel (actually a row of the pixels) to be read-out. The read-out line is driven by an external current source to make the source followers operate correctly. Such pixels are organised in a two-dimensional array and additional access transistors enable to read out one row of such pixels at time.

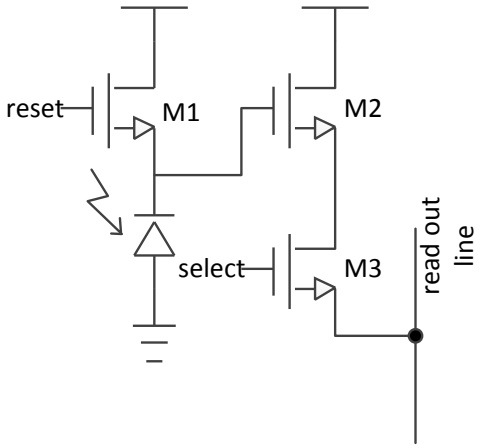


Fig. 1.1: Three transistor active pixel sensor topology.

Today's video cameras vary in number of pixels, quantum efficiency of the photodiodes, fill factor or speed, but the principle of their operation is the same, namely they are all so called frame-based architectures. The key word *frame* means that all these cameras read a series of frames at a constant rate, regardless if the pixels values has change or not. As a result, the frame-based vision sensors produce huge amount of data which contains a lot of redundant information. All pixels hold the same global amplification and the global exposure time. The exposure time is the time how long the charge under the photo diode is accumulated. Hence, such camera is not able to capture image where brightness varies widely at one scene - dynamic range is limited [1]. Most conventional frame-based vision

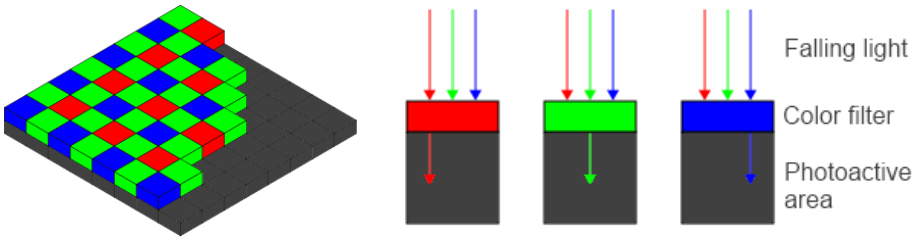


Fig. 1.2: Bayer masks. Different color filters are situated over each single pixel and organised into a color mosaic pattern. Masks are band-pass filters centred a red, green and blue colors (RGB). Our eyes have more sensitivity to green light. For this reason the green filter is placed twice more.

sensors use color filter arrays (*CFA*) to extract color information from the visual scene. Filters are placed over each pixel. Typically they are band-pass filters centred a red, green, and blue colors (RGB). They are usually arranged in groups of four filters (RGBG) called Bayer masks shown in Figure 1.2

[1]There exist frame-based approaches aimed to extend dynamic range, as for instance *dual exposure*, or *multiple exposure*.

1.2 Biological Vision Perception

On the first sight, visual system seems to be uncomplicated, since it is considered to project an image, similarly to the picture obtained by the camera. In fact, the image which reaches our brain, is a result of a complex process, such that we receive only relevant information. *"The most vital part of our vision perception system is the retina, which is the part of the brain, (...) where we begin to put together the relationships between the outside world represented by a visual sense, its physical image (...) and the first neural images"* [1]. The retina is located under eyeball and consists of a thin photosensitive tissue located under eyeball and it converts an optical image into a neural image[2]. There are two types of photoreceptors: rods and cons. The cons are mostly placed in the center of retina, they are responsible for color vision and operate best in bright illumination. Rods are almost entirely responsible for night vision, since they are very sensitive to low illumination of light. High resolution vision is achieved with cones, while rods provide better motion sensing and better sensitivity to low light levels. Very important quality of our vision system is that it has high dynamic range of illumination levels, achieved by horizontal structures located in the lower areas of the retina, called horizontal cells and bipolar cells. Horizontal cells average input from a group of adjacent photo receptors in the eye. In other words they create a diffusive network, which performs averaging of the light level falling on the retina. Then, the bipolar cells can track the difference between the averaged light from the horizontal cells and from a single photoreceptor and get excited basing on this difference. Such solution helps photoreceptors to adapt to different light levels and, as a result, the eye can sense an image with a dynamic range of over 8 decades[3]. Beyond the retina there can be found neuronal structures performing complex image extractions, i.e. cells in VR area of visual cortex, responsible for selective attention [4] cells in V1 area, also located in visual cortex responsive to bars of a particular orientation and direction, cells in MT area sensitive to motion. These structures are based on spike-based nature, i.e. if there is a particular pattern or light intensity variation detected, it will be indicated by a neuronal spiking.

1.3 Frame-Based Camera vs. Retina

The human vision system is much more sophisticated than it might seem and is much more different from a frame-based camera than is evident. It should not be surprising that human's eye outperforms present commercial cameras. High dynamic range is probably the greatest advantage of human eye over commercial cameras. Even though cameras are able to take images of darks scenes, for instance inside a dark room and also take images outdoor in strong light, it is not possible to capture image where brightness varies widely at one scene. The reason is a global adaptation to the light conditions. It can be seen, for instance, while taking picture in a dark room with window, where the outside scene is very bright. Typical cameras are only able to capture details outside the window, while keeping the inside room very dark, or conversely inside room being bright and detailed, while the outside view too bright (overexposed). By contrast, cells in the human retina have local adaptation. As a result, we can see the interior of

the mentioned room, as well as the outside window view. Our adaptation to brightness levels is practically instantaneous due to horizontal connections, performing local area averaging process. What is more, thanks to edge enhancement, our vision system can amplify edges, such that they look more visible to us than in fact they are. It is achieved by illusive colors enhancement of two adjacent areas to increase contrast between them.
Another difference between frame-based cameras and the biological perception system is how much data is extracted from them. Cameras can boast higher and higher resolution nowadays. Unfortunately, with greater number of pixels, it becomes more and more challenging for the camera's bandwidth to transfer and process such huge amount of data from the pixels matrix. The problem is hidden under the frame-based nature of present cameras. Since they simply take series of snapshots at a constant rate, a great number of pixels are sampled repetitively even though their readout values don't change. Interestingly, the human vision system performs much preprocessing already in the eye and applies filtering in order to bring only relevant information to brain. For instance, a spatio-temporal filtering in the eye helps to subtract details about movement and then focus eye's attention on this moving object. Such approach is much more efficient in terms of bandwidth. Finally, electronic circuits dissipate a lot of energy, and very often they need additional cooling system. On the other hand, human body is much more power efficient.

1.4 Event-based Vision Sensors

The advantages listed in the previous section give a good reasoning for a desire to emulate biology's system. There are already implemented bio-inspired vision sensors, which perform local gain control and computation, and generate asynchronous stream of data with only a relevant information regarding the image. They are called Event-Based or Address Event Representation (*AER*) Vision Sensors, because they abandon a frame-based architecture and respond asynchronously in form of events. In other words, each pixel generates events independently represented by their unique address resulting in a stream of pixel address-events on the output[2].
The pioneer of the neuromorphic vision sensors was Carver Mead and Misha Mahowald [5, 6]. Later, Boahen's group in [7] modelled the four predominant ganglion-cell types in the cat retina by microcircuits in silicon. Another significant work was published by Lazzaro in [8], focused on building neuromorphic systems without conventional processor. Later, more and more bio-inspired vision sensors were implemented with better performance and lower mismatch, for instance the first AER vision sensor used in industry [9]. Finally, a famous high dynamic range temporal contrast vision sensor developed by Lichsteiner et al. in [10] started faster development of the bio-inspired vision sensors with a reasonable specification.
In the biological vision system a lot of processing happens in the very beginning of the image sensing. As it was mentioned before, in the retina we can find the cells responsible for local averaging helping to adapt to

[2]More detailed description of the AER operation principle will be described later in next chapter.

different brightness levels. Similarly, based on what kind of processing is performed inside their pixels, most of the AER vision sensors can be divided into two main groups: spatial contrast vision sensors and temporal contrast vision sensors

1.4.1 Spatial Contrast Sensors

Spatial contrast sensors mimic the processing performed by the horizontal diffusive network in the retina. Good example is a work of Misha Mahowald in [11]. She implemented the vision sensor with a resistive horizontal network and pixels operating similarly to the bipolar cells. The pixels compared their inputs with a local averaged value through the resistive network, resulting in contrast enhancement. Alternative implementation of a spatial contrast sensor was proposed by Boahen in [12]. He replaced the linear resistive network with much more practical to implement in CMOS process current mode diffuser network. Further improved version of this vision sensor proposed Leñero-Bardallo et al. in [13, 14]. The latter work showed a dual mode vision sensor with pixels which spike with a frequency proportional to the difference between their photocurrent and the average photocurrent from a 4-pixel neighbourhood. Alternatively pixels could operate as contrast detectors, where pixels with sufficient spatial contrast spiked only once (Time to First Spike).

1.4.2 Temporal Contrast Sensors

The above described spatial contrast sensors produce events continuously when the contrast in the image is detected. Another type of the vision sensors, called Temporal Contrast Sensors, produce events only if the change is detected. As a result they transmit always vital information. Temporal Contrast Sensors mimic rods, which strongly react to a change/movement. Rods are sensitive enough to react to a single photon of light stimuli [15]. Hence, they are capable of attracting our attention to movement even in very dark light conditions. In order to mimic such behaviour pixels compute a *"temporal contrast"*, i.e. they compare their current output with the output from the moment before and respond if these values differ. Since each pixel responds according to the local relative illumination changes, the Temporal Contrast Sensors are independent from the variations of the scene illumination. Lichtsteiner et al in [10] implemented a Dynamic Vision Sensor (*DVS*), which produced events proportional to the rate of the illumination change. Figure 1.3 shows an operation principle of temporal contrast pixel. The pixel generates ON events and OFF events, which represent increase and decrease of the light intensity respectively. It operates as a pulse density modulator, where number of ON and OFF events generated depends on the rate of the change of the illumination. Due to use of logarithmic amplification and differencing circuits, this design can boast of having low mismatch, high dynamic range, good sensitivity and low latency. Additionally, since the sensor operates as a differentiator, only vital information is encoded by the events.

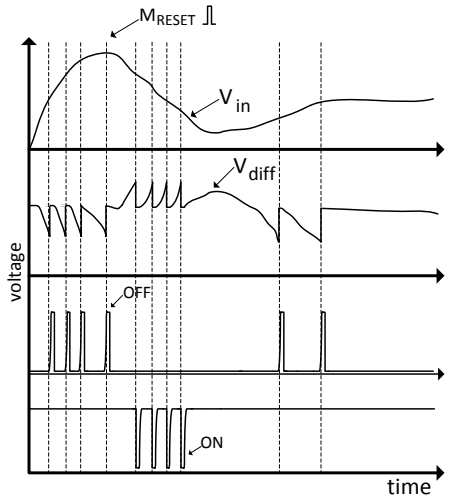


Fig. 1.3: Operation of the single pixel in temporal contrast vision sensor. V_{in} is a transduced photocurrent from the incident light. V_{diff} is a result of the differentiation of V_{in}, meaning it only senses changes. V_{diff} is compared against the positive and the negative threshold by the comparators. If the threshold is exceeded it is indicated by the ON or OFF events (each spike represents one event). A reset signal M_{RESET} balances the differencing amplifier back to the fixed voltage (usually mid-point between power rails).

1.4.3 Event-Based Vision Sensors in Color

Complete image needs to contain information about spectrum of the falling light. Color information is not only vital for a consumer digital cameras industry, but it is critical in many manufacturing processes where robots and automotive devices are involved. For instance, it is essential matching samples of various colored materials or operating mobile equipment for which the operating signals are colored lights [16].
All the spatial contrast and temporal contrast bio-inspired vision sensors described so far are monochromatic. The limitation of the bio-inspired vision sensors to only one color shade is mainly due to their low fill factor and low resolution. Placing discrete color filters (Bayer mask) over bio-inspired pixels to extract color information would further reduce the sensor resolution at least three times. Hence, an alternative method comprised from stacked photodiodes was implemented in bio-inspired vision sensors. The idea of using stacked photodiodes for color filtering was commercialised by company Foveon in 1980. Good example of such sensor is an implemented by Olsson and Häfliger camera using a buried double junction (BDJ) and an integrate-and-fire topology [17]. The information about the color of the incident light is extracted there from the spiking frequencies. In 2011, Leñero-Bardallo et al. published the first tri-color AER vision sensor [18] employing three stacked photodiodes representing color filtered light intensities by a pulse density modulation. Recently this vision sensor was used to indicate the presence of flames by monitoring variations of near infrared radiation [19].
But the neural processing of color is not just based on filtering light at three different wavelengths and representing them by a pulse density modulation. We do have cones with three different pigments that can filter light approximately at red, green, and blue colors, but we do not perceive col-

ors as separate hues [20]. Certain pairs of colors are mutually exclusive. Actually there is an opponency between red and green and between blue and yellow colors. So far there has been implemented vision sensor [21] whose pixels spike when they detect transitions between two colors: red and blue. It emits digital events indicating whether the wavelength of the incident light becomes *bluer* or *redder*, what is equivalent to the red-green opponency.

1.5 Goals and Outline

The overall goal of this thesis was to model the color visual processing of the retina by implementing a temporal contrast vision sensor able to mimic full color opponencies in human eye. The designed sensor should track changes between three primary colors and report whenever the ratios between these three color intensities has changed and became *bluer*, *redder* or *greener*. The reason to mimic the human retina behaviour was to exploit the advantages of the eye listed in the previous sections. This work can be treated as an extension of the design implemented by Berner and Delbrück in [22]. They implemented a single pixel made up by two stacked photo diodes (buried double junction) providing two different color spectra.

The first chapter, *Background*, describes how the light is measured and amplified by electronic circuits, called *photoreceptors*. Later, different approaches and topologies of photoreceptors are explained and their performances are compared. In most details this chapter focuses on the photoreceptor design which was implemented in this work. The rest of this chapter explains the *Address Event Representation* protocol principle and describes its building blocks. The next chapter, *Tri-Color Change Bio-Inspired Sensor - 3C-TVS*, describes the steps taken to design the vision sensor called *3C-TVS*. It explains the principle of operation of a single temporal contrast pixel which employs three stacked photo diodes. This chapter provides also simulation results, design considerations of building blocks, the pixel characterisation, and finally describes layout design of the implemented vision sensor. In chapter *Experimental Results*, a laboratory setup is shown and experimental results are presented. The results have been submitted for publication (two papers available in Appendix E and F). Finally, the work is concluded in the last chapter and future work is discussed. In appendixes additional work is documented, for instance: the developed PCBs, 3D printed custom mounts and the measurement setups. Also, additional photos of the actual chip and figures showing the layout of the implemented sensor are provided.

Chapter 2

Background

This chapter explains how the light is converted into a current and describes how this current is amplified by, so called, logarithmic photoreceptors. Different approaches and topologies of photoreceptors are explained and their performances are compared. In most details this chapter focuses on the photoreceptor design which was implemented in this work. The rest of this chapter explains the *Address Event Representation* protocol principle and describes its building block.

2.1 Photosensors

In order to measure the intensity of the incident light, it has to be converted into different physical form able to be quantified. Sensing light in semiconductor process is realized by the photo active structures able to absorb energy from the falling photons and convert it into a photocurrent. This process is called a photoelectric effect. Figure 2.1 shows a simplified model of a photodiode. The photodiode is formed by a PN junction.

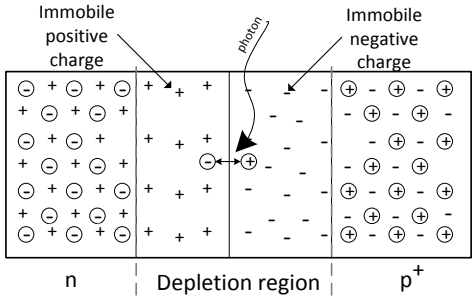


Fig. 2.1: Simplified model of a photodiode and a depletion region formation.

Close to the PN junction mobile charges diffuse and form a depletion region. A falling photon excites an electron from the valence band into the conduction band and forms an electron-hole pair. The energy of one photon is described by [23]:

$$E_{ph} = \frac{hc}{\lambda} \tag{2.1}$$

where h is Plank's constant, c speed of light in vacuum and λ the light wavelength. The minimum energy for photon to interact with silicon is equal 1.1eV, called the silicon band gap. From the Equation 2.1 we can find

the maximum wavelength boundary providing enough energy to interact with silicon. The maximum wavelength is 1125nm [23], what means that the silicon photosensors can detect light within the visual spectra and also near infrared light. Silicon for wavelengths longer than 1125nm is transparent.
The charge created in this way usually recombines immediately, unless there exhibits an electric field. Due to a depletion region the electron-hole pairs created in this way are separated by the electric field resulting in reverse current called photocurrent. The photodiode can be modelled as a current source linear with a incident light intensity. Figure 2.2 shows

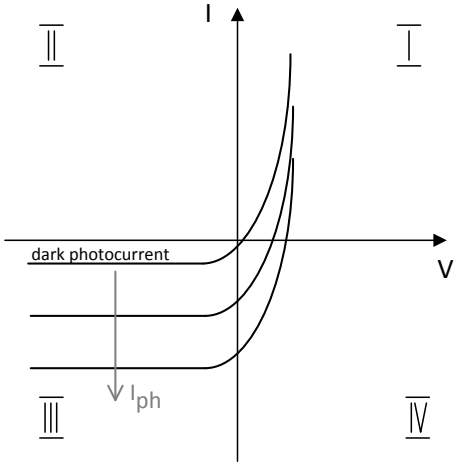


Fig. 2.2: Current-voltage photodiode characteristics.

a current-voltage photodiode characteristics divided into four quadrants. The lower curves show behaviour of the photodiode under different illumination. The photodiode operates usually in quadrant III, because it behaves as a stable current source in this region. The top curve in Figure 2.2 shows a photodiode behaviour with no incident light. Even though there are no photons recombining the photodiode generates a small reverse current, called dark current. The dark current is regarded as a noise because it limits a minimum detectable photocurrent. The values of the dark current in photodiodes is reasonably small due to a depletion region which sufficiently separated charges and limits a random generation of electrons and holes [24] in PN junction.
It should be mentioned that the number of electrons is always lower than the number of incident photons. The measure of efficiency how much charge is collected from the incident photon is called a quantum efficiency. It depends on many semiconductor process parameters, for instance absorption coefficient of the semiconductor, diffusion length, recombination life.

Wavelength-Dependent Absorption Depth

Stacked photodiodes on different depths can be used to extract color information because the absorption coefficient of photons in silicon depends on their wavelength [25]. How deep a photon can penetrate silicon without being absorbed depends on its available power. The photon's power as a

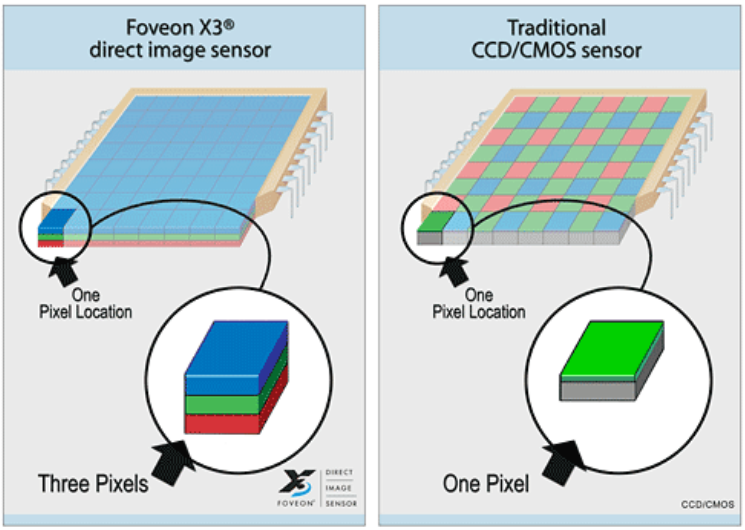


Fig. 2.3: Comparison of a Foveon X3 image sensor (left figure) with a traditional CMOS image sensor (right figure). Figure adapted from [26].

function of a depth x from the surface is [25]:

$$\rho(x) = \rho_0 \cdot e^{-\alpha x} \tag{2.2}$$

ρ_0 is the incident power per unit area of monochromatic light incident on the semiconductor surface and parameter α is an optical absorption coefficient which depends on photon wavelength. Based on this equation there is a higher probability that photons with longer wavelengths penetrate silicon deeper before they are absorbed and conversely, photons with shorter wavelengths get absorbed closer to the surface. Already in 1980 company Foveon commercialised the use of stacked photo diodes for color filtering. Figure 2.3 compares a Foveon X3 image sensor with a traditional CMOS image sensor with a Bayer mask. Area of one pixel in Foveon X3 image sensor is equivalent to three traditional pixels. Detailed structure of a single pixel realised by stacking three photodiodes in a standard 90nm CMOS process is shown in Figure 2.4. The cross section in Figure 2.4a illustrates

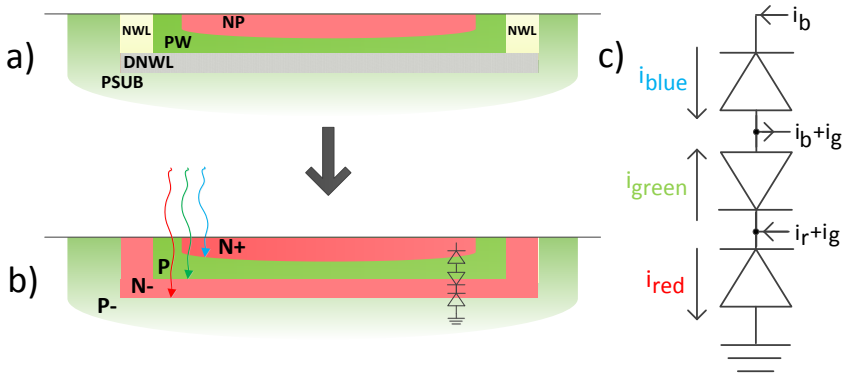


Fig. 2.4: a) Cross section of the stacked photo diodes, b) Formation of PN junctions, c) Equivalent schematic of the stacked photo diodes.

how the stacked photo diodes structure is realized in CMOS process with a deep nwell option. Three PN junctions are formed: N+/PW, PW/DNWL

and DNW/PSUB (Figure 2.4b) . Equivalent schematic for the stacked photodiodes in Figure 2.4c shows that these photo diodes are connected in series. They provide three photo currents. Keeping in mind that photons with shorter wavelengths recombine closer to the surface, these three photo currents can be identified as follows. The top photo diode provides a photo current I_b proportional to blue spectrum, the middle photo diode provides the sum of photo currents $I_b + I_g$ proportional blue and green spectrum and the last photo current $I_r + I_g$ represents green and red spectrum.
The buried double junction (BDJ) mentioned before is also an example of stacked photo diodes but using only two PN junctions instead of three.

2.2 Logarithmic photoreceptors

The photoreceptors are transducing light into an electrical signal. An ideal photoreceptor should have a great dynamic range, stable gain, high sensitivity and high speed for different light conditions. Only a biological photoreceptor can boast of such figures of merit. That is why so called *bio-inspired* vision sensors built in CMOS try to take an inspiration from nature. The basic idea behind the bio-inspired photoreceptor is to use a transistor operating in the subthreshold region. Using a transistor operating in the subthreshold region region, due to its logarithmic I-V dependency, improves a dynamic range of a photoreceptor. Since the light intensity in a typical scene can vary over a wide range, logarithmic compression helps to encapsulate widespread input signal into the smaller output. The performance of the logarithmic photoreceptors will be described from the most basic one, up to the photoreceptor implemented in this work.

2.2.1 Subthreshold Region Operation

Circuits shown in this thesis are based on keeping some transistors in a subthreshold region. The drain current of the transistor in a subthreshold region can be approximated by the EKV model [27]:

$$I_D = I_S \cdot e^{\frac{V_G - V_{T0} - nV_S}{nU_T}} \tag{2.3}$$

Parameters used for modelling the transistor are given in Table. 2.1. The above equation indicates that the drain current in the subthreshold region increases exponentially with V_{GS}.

Parameter	Equation	Parameter name
V_{T0}		the gate threshold voltage
n	$\frac{C_{ox}+C_{j0}}{C_{ox}}$	a subthreshold slope factor
t_{ox}		thickness of the thin oxide under the gate
C_{ox}	$\frac{K_{ox}\epsilon_0}{t_{ox}}$	a gate capacitance per unit area
C_{j0}	$WL\sqrt{q\epsilon_0 N_{sub}/(4\Phi_F)}$	a depletion capacitance
β	$\frac{W}{L}\mu C_{ox}$	
I_S	$2n\beta U_T^2$	specific current
U_T	$\frac{kT}{q}$	thermal voltage
V_G		gate voltage
V_S		source voltage
N_{sub}		channel doping
Φ_F		Fermi potential

Table 2.1: Parameters used further in the thesis.

The subthreshold slope factor n in exponent of Equation 2.3 measures the effect that the source voltage V_S has on the drain current[28]. From Equation 2.3 the transconductance g_m and the conductance g_s (visible from the source) can be derived:

$$g_m = \frac{\partial I_{DS}}{\partial V_g} = \frac{I_{DS}}{nU_T} \tag{2.4}$$

$$g_s = \frac{\partial I_{DS}}{\partial V_s} = \frac{I_{DS}}{U_T} \tag{2.5}$$

2.2.2 Passive Logarithmic Photoreceptors

Figure. 2.5 shows a simple source-follower photoreceptor [29], from group of passive logarithmic photoreceptors. It consists of a single NMOS transistor *M1*. Equation 2.3 for the drain current of the transistor operating in the subthreshold region in now:

$$I_{ph} = I_S \cdot e^{\frac{V_{bias}-V_{T0}-nV_{out}}{nU_T}} \tag{2.6}$$

Equation 2.6 can be rewritten now with respect to the output signal V_{out}.

$$V_{out} = \frac{V_{bias}}{n} - \frac{V_{T0}}{n} - U_T ln(\frac{I_{ph}}{I_S}) \tag{2.7}$$

This expression reveals the logarithmic dependency between the photocurrent I_{ph} and the output voltage V_{out}. For the e-fold change of the light intensity, the voltage at the source of this transistor decreases with a factor of a thermal voltage U_T. Such approach makes possible to compress wide range of a light intensity into output voltage limited only to power rails. A tune-able voltage V_{bias} can be used for adjusting the output DC level of the photoreceptor. A similar configuration of a passive logarithmic photoreceptor is shown in Figure 2.6. By shorting the gate with the source of the transistor M1, V_{bias} is avoided. The appropriate range of operation is set here by choosing the right dimensions of the transistor M1.

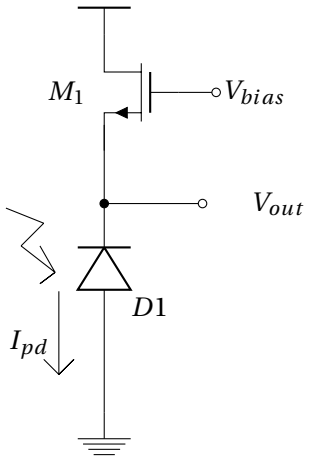


Fig. 2.5: Photoreceptor with an adjustable bias, $I_{ph} - photocurrent$

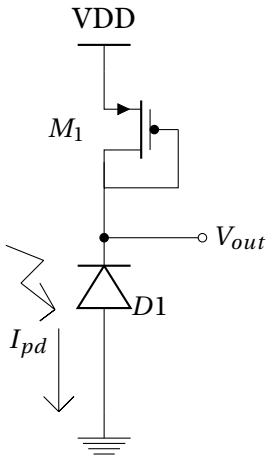


Fig. 2.6: Diode connected logarithmic photoreceptor

If we assume that the transistor M1 always operates in subthreshold region, the Equation 2.3 for the drain current of the transistor operating in the subthreshold region can be rewritten as follows:

$$I_{ph} = I_S \cdot e^{\frac{n \cdot V_{DD} - V_{T0} - V_{out}}{nU_T}} \tag{2.8}$$

Equation 2.8 can be rewritten now with respect to the output signal V_{out}.

$$V_{out} = n \cdot V_{DD} - V_{T0} - U_T \cdot n \cdot ln(\frac{I_{ph}}{I_S}) \tag{2.9}$$

Similarly to the the previous configuration, the photocurrent I_{ph} is logarithmically amplified. This time, the output voltage changes by factor of $U_T \cdot n$ per each e-fold light intensity change.
Both described photoreceptors have a significant drawback. Since the *Thermal Voltage* ($\frac{kT}{q}$) depends on temperature, the gain also depends on temperature changes. A typical gain coming from the sub threshold slope for CMOS is approximately $60mV/decade$ at 273K and will vary with temperature. Additionally, there is a possibility that for high photocurrents I_{ph} transistor M1 will leave a subthreshold region, resulting in gain

distortion.
The gate-source capacitance of transistor M1 is usually much smaller than the junction capacitance of the photodiode: $C_p >> C_{gs}$. Hence, the source follower photoreceptor shown in Figure 2.5 can be simplified in a small signal analysis to a first order system consisting of a source conductance g_s in parallel with a reverse biased diode capacitance C_P - Figure 2.7. Using

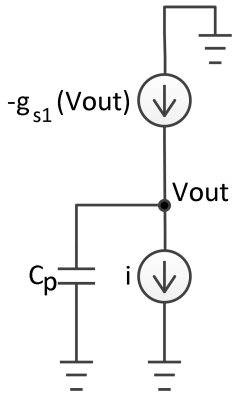


Fig. 2.7: Small signal model of a Photoreceptor from Figure 2.5

Equation 2.5, the nodal analysis gives a time constant:

$$\tau_1 = \frac{C_P}{g_s} = \frac{C_P U_T}{I_{ph}} \tag{2.10}$$

It means that the bandwidth of this circuit is inversely proportional to the background photocurrent. For a small photocurrents (low illumination) it takes too much time to charge the relatively big capacitance of the photodiode.

2.2.3 Photoreceptors With Negative feedback

So called transimpedance amplifier shown in Figure 2.8 [10] has more stable and robust gain thanks to a beneficial effect of a negative feedback. The photoreceptor in Figure 2.8 is a modified version of the adaptive logarithmic photoreceptor originally described in [29] by Delbrück and Mead. The difference here is a resignation from the capacitive feedback which gives higher gain for transient stimuli described as adaptiveness. Resignation from these capacitors however decreases the area needed to implement the pixel and, hence, improves a fill factor of the pixels array. Similarly to the previous circuits, it converts a photocurrent into voltage logarithmically. The input stage is realized in the same way as in Figure 2.5, using transistor M1 as a source-follower sourced by a photodiode D1. The output voltage, V_p of the input stage is amplified by a common source amplifier (transistors M2 and M3), of amplification equal A_{amp}. Finally, the amplified and inverted voltage V_{out} is fed back to the gate of transistor M1. Additionally, the voltage V_{bias} can be used to adjust the gain of the common source amplifier. Figure 2.9 gives a conceptual understanding of the circuit. The photocurrent is considered (from a small signal point of view) as consisting of the background constant current I_B (setting the bias conditions) and a small changing signal i. Recalling Equation 2.5, for a

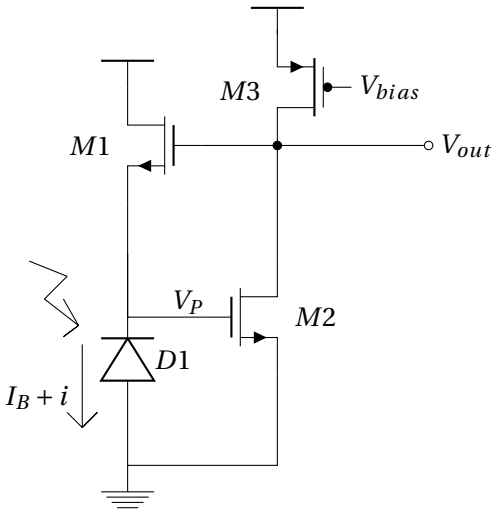


Fig. 2.8: Photoreceptor with negative feedback

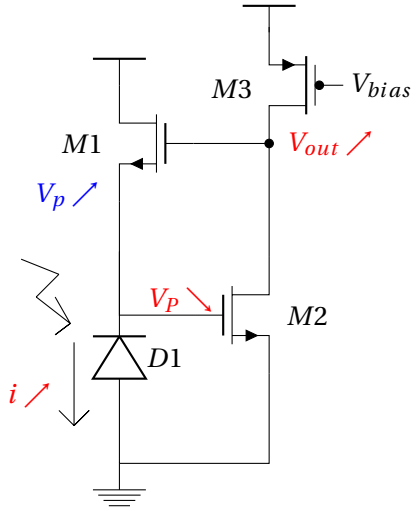


Fig. 2.9: Conceptual explanation of the photoreceptor from Figure 2.8.

small increase of the photocurrent i, the source of transistor M1, V_p will change by U_T. This gives a gain of the first stage equal *Thermal Voltage* - U_T. Voltage V_{out} is amplified $-A_{amp}$ times and fed back to transistor M1. Since the common source amplifier has much higher gain, the gate voltage of transistor M1 increases, forcing V_P to raise, instead of dropping as indicated by a blue label V_P. The negative feedback can be simplified to the inverter with gain $-A$, as shown in Figure 2.10. Thanks to this solutions the source of transistor M1 does not hit the rails as it was in the Passive Photoreceptors and transistor M1 remains longer in subthreshold region, resulting in bigger dynamic range. This circuit is much more useful to model a biological system, because it adapts to the light intensity similarly as a photoreceptor in human eye. For low frequency signals the amplification can be calculated as follows. Recalling Equation 2.3 and Equation 2.7, the gain coming from transistor M1 is U_T. Assuming the gain of the common source amplifier to be infinite, the gate of the transistor M1

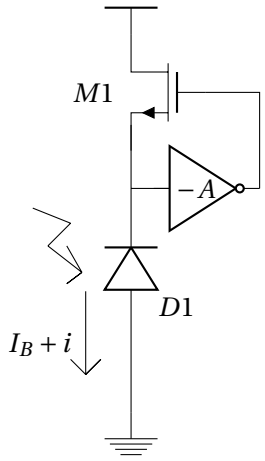


Fig. 2.10: Abstracted schematic of photoreceptor from Figure 2.8

has to change $U_T \cdot n$ to follow the e-fold change of photocurrent, resulting in the low frequency gain [29]:

$$A_{low\_f} = U_T \cdot n \tag{2.11}$$

Please note that the gain is not linear, as the Equation 2.11 suggests, because we specify the gain per the e-fold change of the photocurrent. In contrast, the change in output voltage in response to a linear change in the input photocurrent is:

$$\Delta V_{out} = U_T \cdot n \cdot ln(\frac{\Delta i}{I_S}) \tag{2.12}$$

I_S is a specific current of the transistor operating in subthreshold region.

2.2.4 Small Signal Analysis

The small-signal model of the Photoreceptor with Negative Feedback is shown in Figure 2.11. The small signal model and analysis is based on [30]. According to [30], the following time constants can be recognised :

$$\tau_p = \frac{C_j}{g_{m1}} \tag{2.13}$$

$$\tau_o = \frac{C_o}{g_{m2}} \tag{2.14}$$

$$\tau_m = \frac{C_m}{g_{m2}} \tag{2.15}$$

$$\tau_n = \frac{C_m}{g_{m1}} \tag{2.16}$$

The common source amplifier gain is:

$$A = \frac{g_{m2}}{g_{m2} + g_{m3}}. \tag{2.17}$$

Assuming infinitely high amplifier gain the nodal analysis conducted in [30] gives a simplified low-pass transfer function:

$$\frac{v_{out}/U_T}{i/I_b} = \frac{-n}{1 + s\tau_n} \tag{2.18}$$

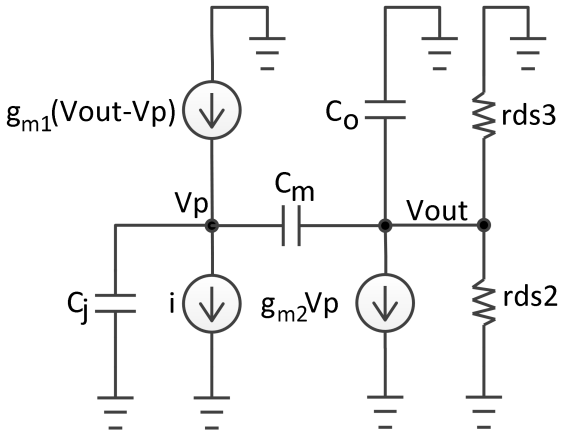


Fig. 2.11: Small-signal model of the Photoreceptor With Negative feedback. C_m - Miller capacitance of the gate-source capacitance of transistor M1 and gate-drain capacitance of transistor M2, C_o - output capacitance consisting of source-bulk capacitance of transistor M2 and drain-bulk capacitance of transistor M3, C_j - capacitance of the depletion region of the photodiode, g_{m1} - transconductance of transistor in subthreshold region as in Equation 2.4

Time constant coming from the diode is inversely proportional to the loop gain, so it becomes very small and can be omitted. The dominant time constant comes from the Miller capacitance:

$$\tau_n = \frac{C_m}{g_{m1}} \tag{2.19}$$

Equation 2.19 shows that the increase of the amplifier gain will not improve the bandwidth. Photoreceptor with negative feedback and with a cascode transistor M3 shown in Figure 2.12 can further improve the bandwidth. The reason of adding the cascode transistor M3 is to further decrease the

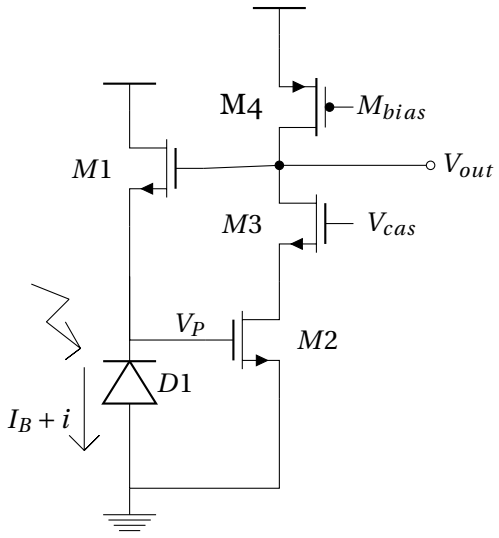


Fig. 2.12: Photoreceptor with negative feedback, with a cascode transistor.

dominant time constant thanks to limiting the Miller effect on the gate-drain parasitic capacitance of transistor M2. Based on a small signal

analysis shown in Figure 2.13 and assuming that transistors M2 and M3 introduce the same resistance r_{ds} Equation 2.18 becomes:

$$\frac{v_{out}/U_T}{i/I_b} = \frac{-n}{1+s\frac{\tau_n}{2}} \tag{2.20}$$

This indicates that adding a cascode transistor improves a speed of the photoreceptor twice. Additionally the gain of the common source amplifier increases what results in further increase of speed (by further decreasing of voltage swing at V_p). When compared with Passive Photoreceptors, the

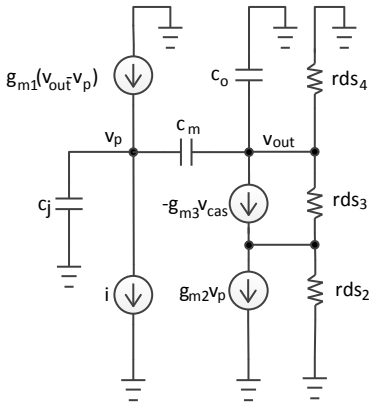


Fig. 2.13: Small signal circuit of a Photoreceptor with Negative Feedback, and with a cascode transistor.

bias current of this photoreceptor is higher then the photocurrent and it is more suitable for driving the capacitive outputs.
In fact, the gain of this amplifier is lower, because the negative feedback cancels any change of the input. However, the more important is the fact that because of the negative feedback, the gain of the amplifier is more stable. Since the next stage serves high-gain amplification, much more important is the linearity of the front-end stage. The Passive Logarithmic Amplifier can easily *hit the rails* of a power supply, what results in bigger distortion. The Photoreceptor with Negative Feedback stays in linear region for much wider range of intensities and has wider bandwidth as it was shown in Equation 2.18. The simulation of the bandwidth of this photoreceptor is included in Chapter *Tri-Color Change Bio-Inspired Sensor*. From the vision sensor design perspective, only stability of the gain of the front-end plays a crucial role. Any nonidealities on this stage will be amplified with high gain resulting in fixed-pattern noise (FPN), i.e. different photoreceptors having different output signal levels for the same illumination. Naturally, because of the process variations there will DC mismatch in V_{out}. However, it will be removed by differencing circuit, which amplifies only a changes of V_{out}.

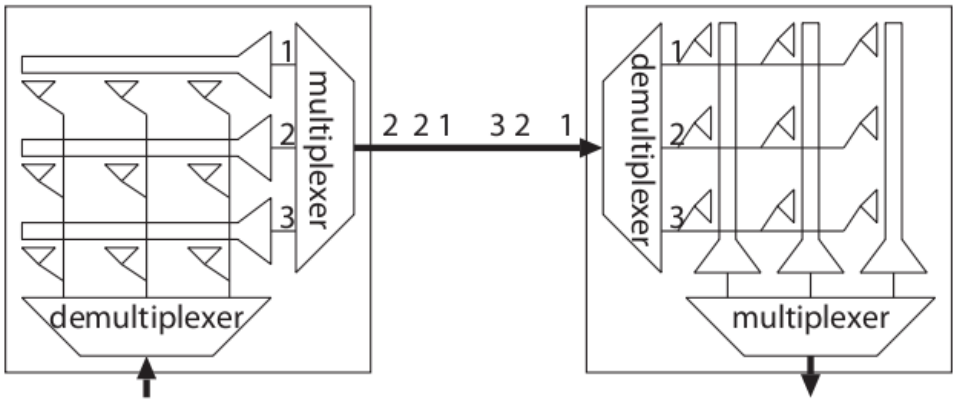


Fig. 2.14: The basic idea of AER Protocol[32].

2.3 AER Protocol

The neural communication is based on a point-to-point communication, what leads to an extensive branching between neurons. Human brain has roughly 10^{11} neurons and each neuron has around 10^4 connections to others [31]. The most convenient and *natural* would be to implement this approach also in so called bio-inspired electronic circuits. However, giving to each neuron an individual connection to the other neurons, such that each *'wire'* is unique and dedicated for every single neuron is impossible from the PCB and ASIC design point of view. Even most advances chips are far from providing such dense wiring. Electronic boards, because of their two dimensional layout organisation are impossible to provide such staggering number of connections. An elegant solution of mimicking such dense communication in the asynchronous nervous system is to use the Address Event Representation protocol *(AER)*. The basic idea of AER is shown in Figure 2.14 [32]. Instead of implementing physically all mentioned connections, the AER Protocol takes advantage of speed of digital circuits and imitates densely wired neural networks. Since the output information of each neuron is represented by its spikes, the AER protocol assigns each neuron its unique address. Any time a particular neuron spikes, it is represented in the AER bus by his address. In Figure 2.14 [32] the AER protocol is used to realise a communication between two two-dimensional neural networks. The left block can be treated as a sender, the right one as a receiver. The neurons send a request to the bus if they generate an action potential. If they obtain acknowledge, their addresses are send through the bus. The final result is a stream of addresses sent to the bus. Even though the transmission of the addresses through the bus is realised serially, statistically the communication should not suffer from delays. That is because the AER communication is much faster than the neuron activity[1]. For a high enough speed the AER protocol gives an impression of a real-time communication.

[1] *"Refractoriness limits the number of action potentials that a neuron can produce per unit time ..."*[33, p.56]

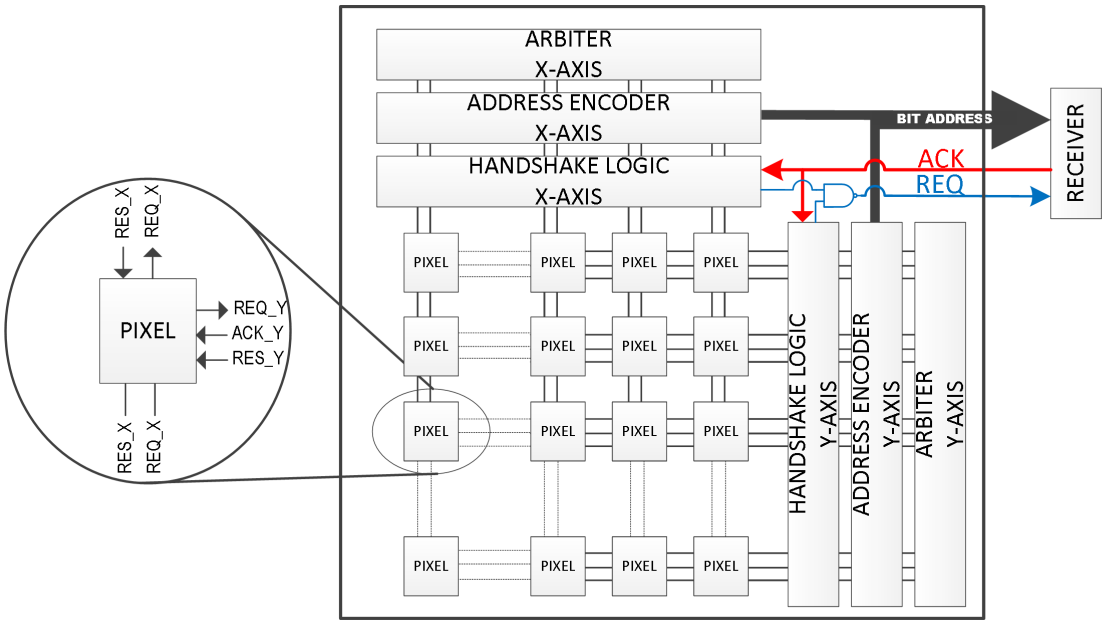


Fig. 2.15: Block level view of the pixel array and the AER communication periphery. The number of pixels is decreased for clarity. Each pixel has 3 horizontal and 2 vertical signals. Handshake logic blocks handle the communication between the array and the AER periphery. Address encoders generate bit address of the pixel which has generated a spike.

2.3.1 Retina AER Communication Protocol

Retina AER communication protocol is a solution implemented in an event-based asynchronous bio-inspired vision sensor. The array of pixels is in fact a two dimensional matrix of cells (rods and cons in the retina). Pixels are spike-generators, whose output (action potential - AP, alternatively called *event*) needs to be read out and sent through the bus to the receiver. The more detailed description of this structure is given in the Figure 2.15, which shows a block level view of the pixel array combined with AER periphery implemented in this work. The vision sensor embedded with the AER communication periphery can be treated as a sender - left block in Figure 2.14. Each pixels shares connections with pixels in the same row and column.

2.3.2 Arbitration

Matrix of pixels create a multiple point-to-point system, where only one bus is shared. It becomes clear that it is possible, that two or more pixels generate an event at the same time. As a result the generated bus address can be lost or become wrong. In order to avoid such problems arbitration needs to be implemented . Whenever one pixel generates AP

(event), first it has to generate a request to the bus for transmission. After the acknowledge, the address of this pixel is exposed on the bus. The arbitration implemented in this work is a full arbitration with a 'Greedy' Arbiter [32] shown in Figure 2.16.

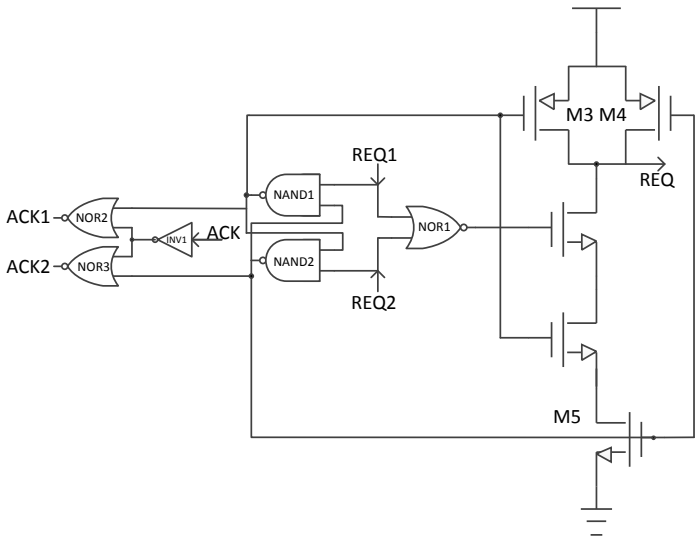


Fig. 2.16: The glitch-free 'greedy' arbiter.

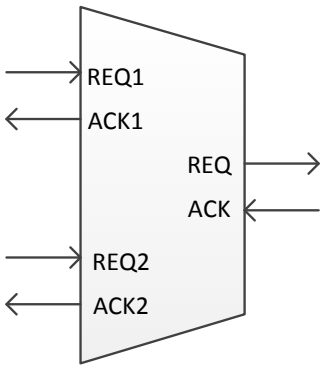


Fig. 2.17: Symbol of the corresponding arbiter.

The basic idea of this arbiter is to use a RS flipflop with active high inputs - NAND1 and NAND2 in Figure 2.16. If there is no request, RS flipflop operates in its restricted combination, where both outputs are high. If REQ1 signal comes low first, the higher input of NAND2 becomes low what blocks signal REQ2. REQ1 signal is the winner and the output REQ signal is generated to the next level. Additional circuit similar to a NAND gate (transistors M1-M5) is used to prevent glitches described in [32]. The arbiters are organised into a binary tree in order to provide a full arbitration, as it is shown in Figure 2.18. The last arbiter has its ACK and REQ ports shorted. In order to provide arbitration of 8 signals, 3 levels of arbiters are used. For instance, if the REQ4 obtains acknowledge, it means that it won over REQ5 in the first level, with REQ6 and REQ7 in the next level and REQ0, REQ1, REQ2 and REQ3 in the last level. The main disadvantage of such arbitration is the greediness. When looking on the table of truth for RS flipflop, the arbiter which grants acknowledge to the one of the requests, for instance REQ1, will then grant the acknowledge for the second request REQ2. As a result, rows which are most active can be favoured.

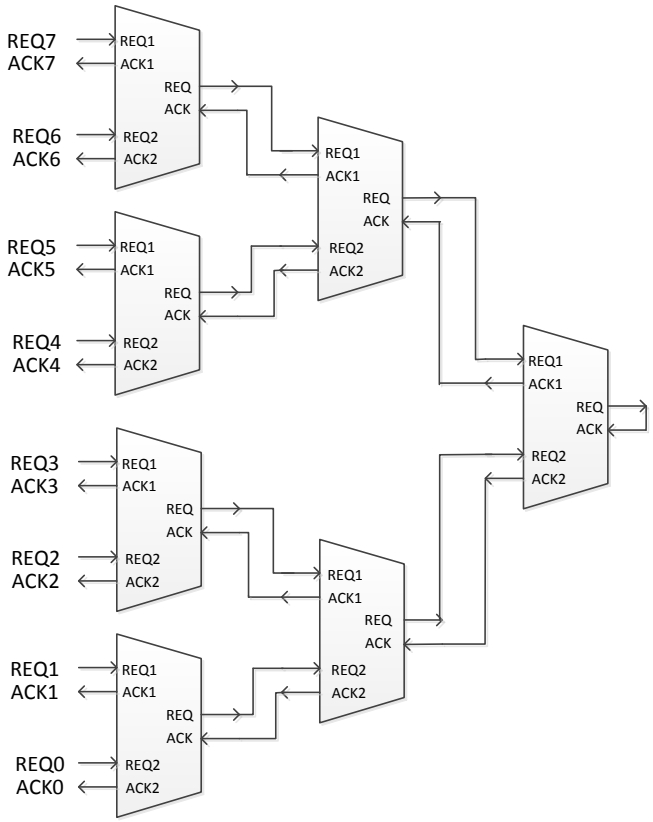


Fig. 2.18: A binary tree of arbiters, handling 8 requests.

2.3.3 Handshake

The handshake circuitry used in this thesis is based on the PhD work of Philipp Häfliger [34]. This is an improved version of the topology implemented by Mahowald in her PhD work [5]. The handshake logic passes through internal signals between the pixel array and arbiters. The organisation of the matrix and the handshake logic is shown in Figure 2.19. The pixel spike initiates the communication procedure by pulling down the REQ_Y signal along the entire row. If the acknowledge is granted by signal ACK_Y, the column request is generated and a bus request is generated by REQ signal. Signal REQ is illustrated by a blue wire in Figure 2.20. After the acknowledge signal comes, the pixel address is read-out and the pixel is reset. It means that the handshake logic handles the external communication, such that it ensures that address of the currently handled pixel is correctly read-out. In Häfliger's approach the handshaking is speed up thanks to latching and pipeline operations in the row level. Namely, in Mahowald's approach the handshake and arbitration of the next pixel is performed no sooner than the previous pixel is handled. In Häfliger's approach the internal handshake can be performed simultaneously during the read-out. As a result. the address of the next pixel is prepared to be sent before the previous pixel is handled. This makes a delay between next addresses read-out shorter and improves a speed.

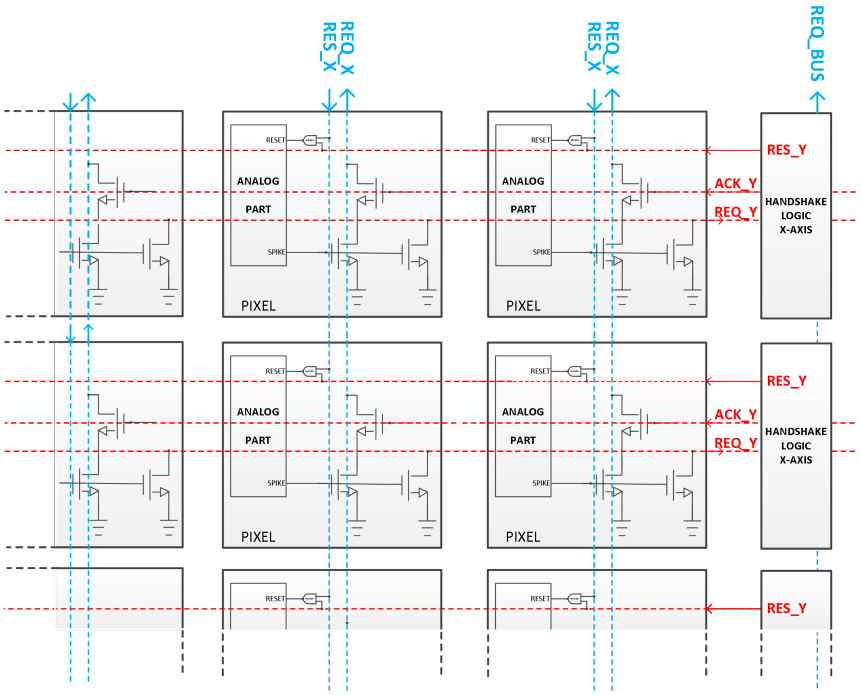


Fig. 2.19: Integration of the handshaking logic and pixels along the x-axis. Also the internal communication circuitry of pixels is shown.

2.3.4 Address encoders

Address encoders are generating appropriate addresses of the currently acknowledged pixel. The principle of their operation is explained in Figure 2.21. Two types of address encoders can be distinguished: *bit 0* and *bit 1*. *Bit 0* consists of a single NMOS transistor which acts as a pull-down and *Bit 1* consists of a single PMOS transistor acting as a pull-up resistor. The designed vision sensor in this work needs 10 bits to encode pixel addresses, e.g. 5 bits are assigned to the row addresses and another 5 bits to the column addresses.

2.3.5 Pixel case

The AER communication protocol with basic blocks was shown in this section. Pixel designed in this work generates 4 independent outputs. Each output needs an unique address assigned. The communication protocol shown in this section is still valid if 4 pixel cells shown in Figure 2.15 are treated as a one pixel. This is illustrated in Figure 2.22. The pixel can be divided into upper and lower part. The lower part has two outputs: ON_down, OFF_down and the upper part: ON, OFF. The only difference when compared with the single output case is that the NAND gate responsible for balancing the pixel after successful address read-out is replaced by the circuit shown in Figure 2.23a. The table of truth shows that the pixel is in reset mode for $(RES\_Y \&\& (RES\_X\_ON || RES\_X\_OFF)) == 1$ [2]. The pixel is balanced for active high RESET. V_refr is a global control voltage adjusting a conductivity of a pull-up transistor, and hence a time needed for a RESET signal to return to VDD after resting a pixel. The less

[2] && is logic OR, || - logic AND

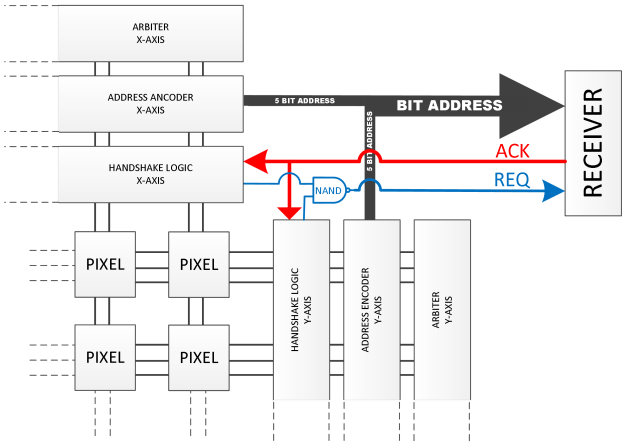


Fig. 2.20: An external handshake communication

conductive transistor is, the more time is needed resulting in longer time of a reset mode.

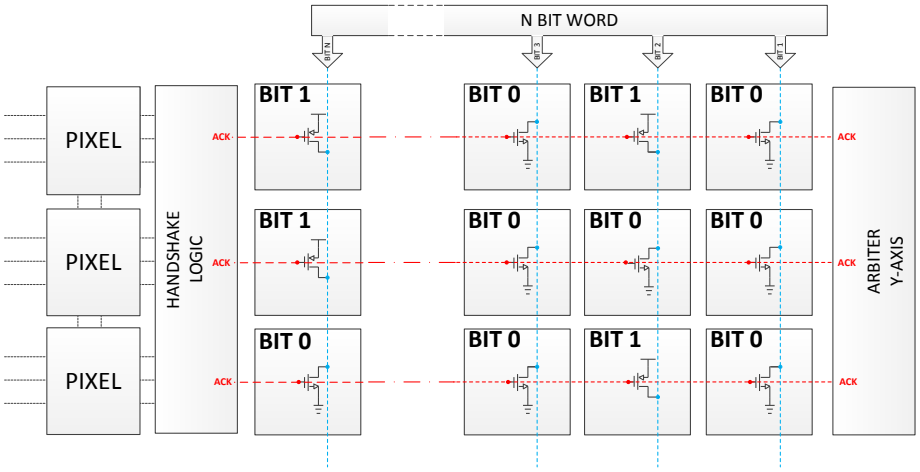


Fig. 2.21: Block diagram showing the principle of address encoders. Address encoders can be distinguished into BIT0 and BIT1. Each row has different configuration of these bits. When one row receives an acknowledge, signal ACK becomes high. As a result, BIT0 blocks pull down and BIT1 pull up vertical lines. These lines represent an N-bit word. Another signals were omitted for clarity.

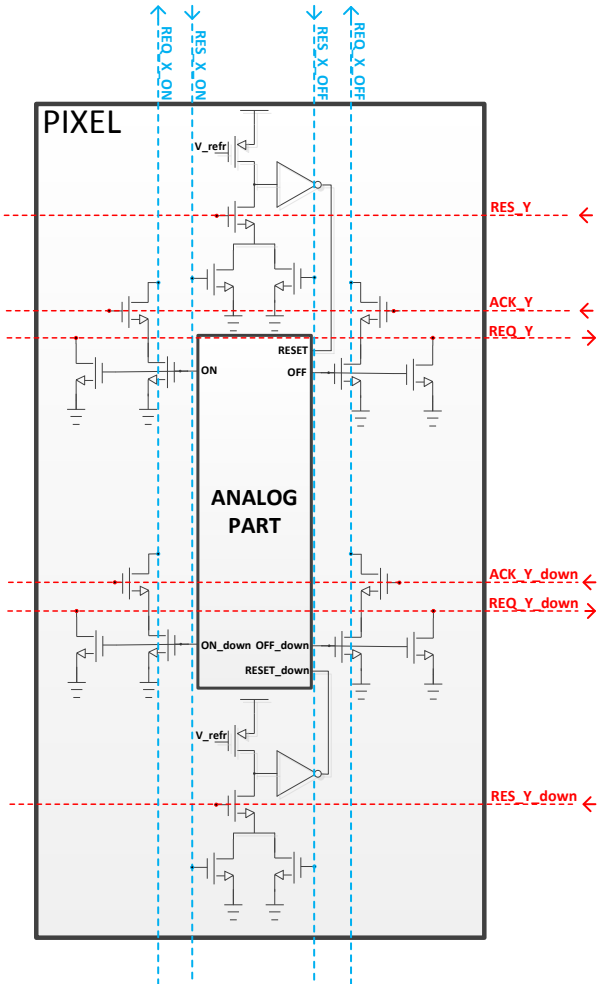
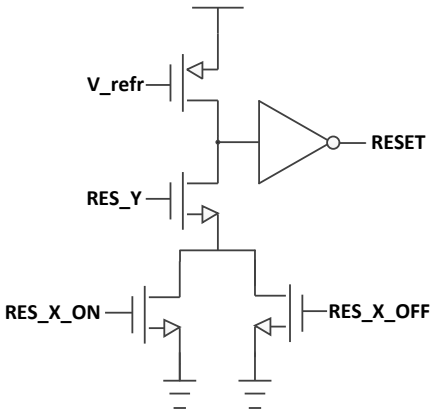


Fig. 2.22: Pixel handshaking internal circuitry. Pixel generates 4 outputs, which need a unique address: ON, ON_down, OFF, OFF_down. NAND gate responsible for balancing the pixel in Figure 2.20 is replaced by a 6 transistor circuit shown in Figure 2.23a



(a) V_ref controls a conductivity of a pull-up transistor, it is used to control how long the pixel is reset.

RES_Y	RES_X_ON	RES_X_OFF	RESET
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	0
1	0	0	0
1	0	1	1
1	1	0	1
1	1	1	1

(b) Table of truth for the balancing circuit. Reset mode is indicated by a red color.

Fig. 2.23: Balancing circuit for a 4 output pixel case

Chapter 3

Tri-Color Change Bio-Inspired Sensor - 3C-TVS

3.1 3C-Pixel

This chapter describes a design and principle of operation of a single temporal contrast pixel which employs three stacked photo diodes. Further, this pixel is called *3C-Pixel* (where 3C stands for 3 colors). Also, this chapter provides simulation results and design considerations of building blocks. Finally whole pixel is tested and characterised by a Monte-Carlo analysis. Second part of this chapter focuses on the entire vision sensor, called 3C-TVS (3-Color Temporal Vision Sensor).

3.1.1 The 3C-Pixel block diagram

The 3C-Pixel can detect changes in color contrast. The output binary event rates indicate the increase/decrease in spectra ratios. Figure 3.1 shows a block diagram of one 3C-Pixel. The pixel topology can be treated as an extension of the design implemented by Berner and Delbrück in [22]. Their pixel comprises two stacked photo diodes (buried double junction) providing two different color spectra, whereas the pixel implemented here is made up by 3 different photo diodes stacked at different depths, providing three different color spectra [18]. The first stage of the 3C-Pixel are the three stacked photo diodes, which provide three photocurrents: I_b, $I_b + I_g$ and $I_r + I_g$. Letter 'b' stands for blue, 'g' - green, 'r' - red. These photo currents are the result of the different absorption spectra of the stacked photo diodes which depend on their depths in the substrate. Photoreceptors convert logarithmically photo currents into voltages V_{ox}. These photoreceptors have high dynamic range and high bandwidth. The photo receptor output V_{ox} is then buffered by a source follower to the second stage - a differencing amplifier. The reason of using the source follower is to isolate the differencing amplifier from the photoreceptor. The second stage amplifiers can produce high frequency noise during a reset mode, which can affect very sensitive output of the photoreceptor. The differencing amplifiers amplify only changes in the ratio of the photocurrents:$(I_b)/(I_b + I_g)$ and $(I_b + I_g)/(I_g + I_r)$. The gains of the differencing amplifiers are determined by the ratio between the input

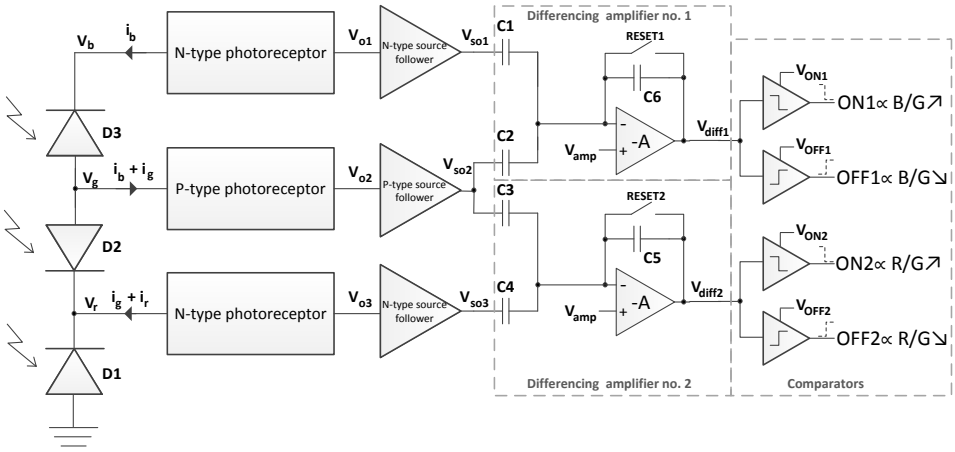


Fig. 3.1: Block diagram of the 3C-Pixel. Capacitors C1-C4 are part of the differencing amplifiers.

and the feedback capacitors. As a result outputs V_{diffx} indicate if the ratios $(I_b)/(I_b + I_g)$ and $(I_b + I_g)/(I_g + I_r)$ have increased whether decreased. If the light level of the pixel changes without a change in the spectrum, the pixel will not react because the ratio remains constant. Thus, the pixel is independent from the global illumination level of the input. If for instance $(I_b)/(I_b + I_g)$ ratio increases, then V_{diff1} drops, and if V_{diff1} exceeds the adjustable threshold V_{ON1}, the pixel produces positive event ON1 [1]. Conversely, if the ratio decreases, the pixel produces OFF1 event. Similarly changes in the second ratio are indicated by ON2 and OFF2 events. After event is generated the differencing amplifier is reset by RESETx signal and the process repeats.
The reason of providing N-type and P-type photoreceptors is the fact that the direction of the middle photocurrent is opposite to the two remaining. The currents I_b and $I_r + I_g$ are drawn from photoreceptors, when $I_b + I_g$ is supplied to the photoreceptor. This yields a need to implement a complementary circuit working identically as the N-type circuit, but drawing an input photocurrent. The reason of implementing a P-type source follower is a linearity, discussed further.

3.1.2 Modelling Stacked Photo Diodes

Stacked diode consists of three different photodiodes stacked at different depths as it is shown in Figure 3.2. Three junction capacitances of diodes can be distinguished: N+/PW, PW/DNWL and DNW/PSUB. As it was shown in Chapter *Background*, the diode capacitance affects the speed of the photoreceptor. Hence, when simulating a Photoreceptor performance, the junction capacitance value should be assumed to be the worst case. The junction resistance is not critical, because for a reverse biased diode, it is infinitely big. Table 3.1 compares the theoretical approximated junction capacitances per area. It also shows estimates of junction capacitances for a stacked diode of size $50\mu x50\mu$. These values are very approximated, because only a horizontal area

[1]The differencing amplifier is inverting, hence decrease in the ratio results in drop of V_{diffx}

Fig. 3.2: a) A cross section of stacked photodiodes

Junction	$C_j[mF/m^2]$	area $[\mu m^2]$	$C_{jo}[pF]$
N+/PW	2.09	2106	4.4
PW/DNWL	0.76	2401	1.82
DNW/PSUB	0.137	2714	0.371

Table 3.1: C_j - junction capacitance

was taken into consideration. Additionaly, the junction capacitance C_j is an $0V$ bias capacitance (only for built-in voltage of junction). In the real application these diodes are reverse biased, what leads to the increase of the depletion region and hence decrease of the junction capacitance. In order to include the influence of the photodiodes in the simulations an N+/PW diode D1 of size $49\mu x49\mu$ is added in parallel to the source emulating photocurrent in Figure 3.3. This type of the diode has the highest capacitance expected from the N+/PW junction. The size $49\mu x49\mu$ is an initially assumed size of the stacked photodiodes used in the pixel design. The photodiode is very big, however based on results in [35], for such area, photocurrent values should be reasonably high. The typical photocurrents are roughly around $10pA - 1n$ the order of magnitude.

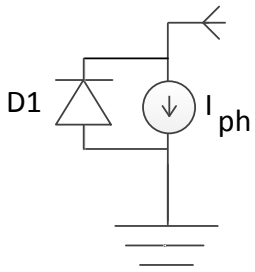


Fig. 3.3: Assumed model of the photodiode used in further simulations.

3.1.3 Front-end Photoreceptor

The front-end photoreceptor is the photoreceptor with the negative feedback and with the cascode transistor described in chapter *Background* (Figure 3.4) [10]. There are two ground reference voltages: 0V and 0.8V. The size of the transistors are collected in Table 3.2.

Transistor	W [μm]	L [μm]
M1	0.4	2
M2	1	0.5
M3	2	0.5
M4	0.5	5

Table 3.2: Sizes of the transistors of the photoreceptor from Figure 3.4

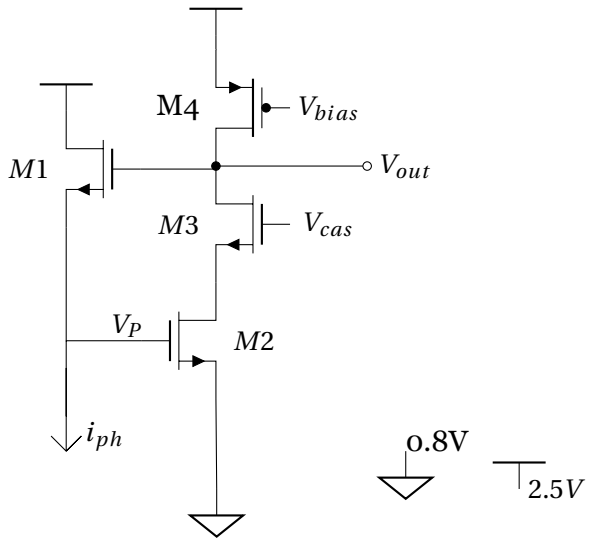


Fig. 3.4: Schematic of the N-type photoreceptor. $V_{bias} = 1.7V, V_{cas} = 1.6V$

The N-type Photoreceptor supplies a photocurrent to the photodiodes, that is consistent with the current direction I_b and $I_r + I_g$ in Figure 3.1. However, the direction of the middle photocurrent $I_b + I_g$ is opposite to the two remaining. This yields a need to design a photoreceptor able to sink a photocurrent.

Transistor	W [μm]	L [μm]
M1	0.4	0.5
M2	1	0.4
M3	2	0.4
M4	0.5	5.7

Table 3.3: Sizes of the transistors of the Photoreceptor from Figure 3.5

Probably the easiest possible solution would be simply to invert the current $I_b + I_g$ by a current mirror. However mismatch of the drain currents in the current mirrors is inversely proportional to the current [27]. Since the photocurrents are of magnitude below nanoampers, the mismatch coming from the current mirror would be unacceptable. Much better solution is to implement a complementary circuit working identically as the one shown in Figure 3.4, but now drawing an input photocurrent. This is achieved by replacing all NMOS transistors by PMOS and vice versa. The resulting complementary Photoreceptor shown in Figure 3.5 works similarly as the N-type photoreceptor, but with an opposite sign in gain. N-type and P-type Photoreceptors were used to sense the three different photocurrents, as it

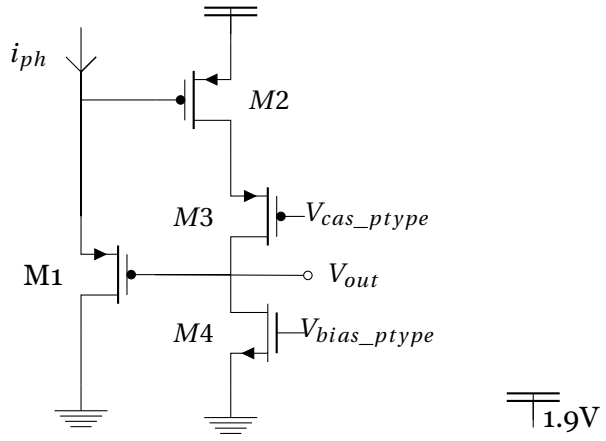


Fig. 3.5: Schematic of a P-type Photoreceptor. $V_{bias\_ptype} = 0.8V$, $V_{cas\_ptype} = 1V$. Sizes of the transistors collected in Table 3.3

is shown in Figure 3.6.

Gain compensation

In order to measure the gain of both photoreceptors, the photoccurent of the diode D2 was swept while keeping other diodes photocurrents constant. Since the ratio of the input photocurrents was kept constant, the sum of the voltages was expected to be flat. The resulting output voltages V_{o2} and V_{o3} are plotted in Figure 3.24a. Since the photoreceptor has a logarithmic I-V dependency, the output voltages are straight lines for the logarithmic X-axis. V_{o3} is close to 2V and V_{o2} close 1V in order to provide proper biasing of the photodiodes, discussed further in this section. The slope of V_{03} is steeper, what results in a non-flat sum of these voltages illustrated by the blue curves in Figure 3.24b, meaning that there is a gain mismatch between these photoreceptors.
In [30] the problem with the gain mismatch between the complementary circuits was solved by manipulating the gain of the second stage amplification. In this work I used a weak dependency between the size of the transistor in subthreshold region and a slope factor. The subthreshold slope factor measures the effect that the gate-source voltage V_{GS} has on the drain current[28]. The subthreshold slope factor equation from Table 2.1 is:

$$n = (C_{ox} + C_{j0})/C_{ox} \tag{3.1}$$

Assuming the gate capacitance be constant, differences in the depletion capacitance affects the slope factor. The depletion capacitance is[36]:

$$C_{j0} = WL\sqrt{q\epsilon_0 N_{sub}/(4\Phi_F)} \tag{3.2}$$

In the N-type Photoreceptor an NMOS transistor is kept in subthreshold region, while in P-type Photoreceptor this is a PMOS transistor. Typically, the NWELL has higher doping level N_{sub} then the substrate - PSUB. As a result the depletion capacitance C_{j0} for the PMOS transistor is bigger. In order to compensate it, the length of transistor M1 in Figure 3.5 has been decreased from 2000nm to 500nm. The resulting sum output voltages after this modification is illustrated by the blue curves in Figure 3.24b. It should

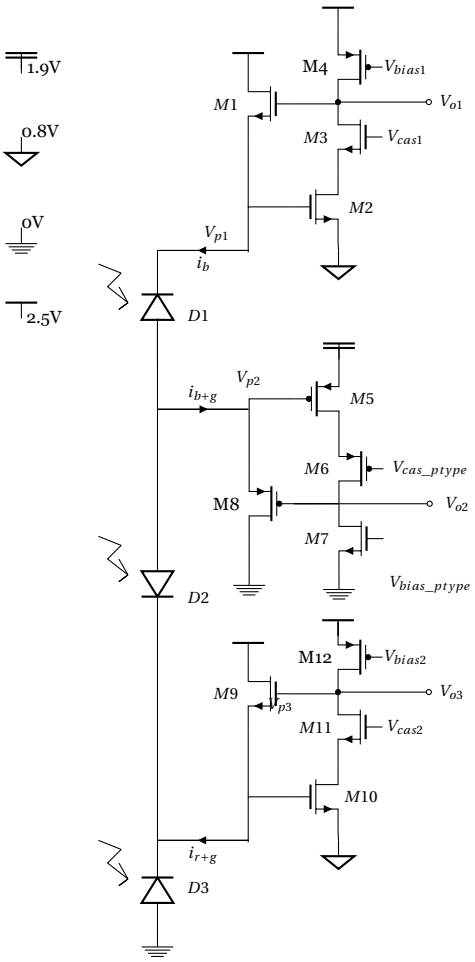


Fig. 3.6: Circuit of the 3C-pixel logarithmic photoreceptors amplifying three different photocurrents.

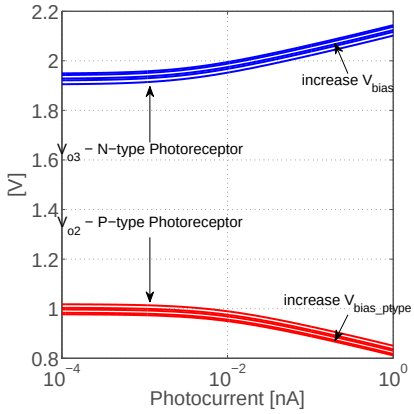
be mentioned that the equation for the subthreshold slope factor shown in Equation 3.1 is only an approximation. In fact the subthreshold slope factor is also a function of the gate voltage[28]. However the simulation results suggest that after this modification the mismatch is canceled.

Bandwidth

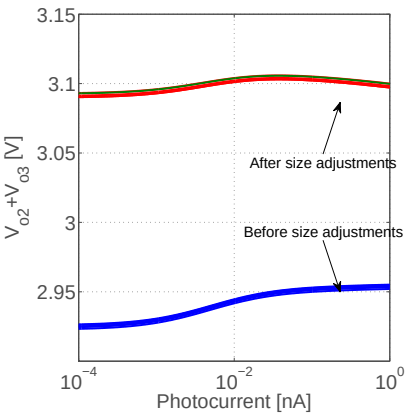
The design considerations regarding the bandwidth refer to both N-type and P-type Photoreceptors.
In chapter *Background* it was shown that the dominant time constant comes from the Miller capacitance C_m and the transconductance of transistor in subthreshold region - g_m:

$$\tau_n = \frac{C_m}{g_{m1}} \tag{3.3}$$

This capacitance is C_{gs} - the gate-source capacitance of the transistor M1, as shown in Figure 3.8. The capacitance C_{gs} couples the output of the common source amplifier with the input of the Photoreceptor. Since the capacitor node from the input side has to follow big changes of the output voltage, the capacitance visible from the input side becomes $A(1+C)$, where



(a) Response of the N-type photoreceptor - blue colour, and the P-type photoreceptor - red colour for different bias voltages



(b) Sum of V_{o2} and V_{o3}, before the size adjustments - blue color, after the size adjustments - red color.

Fig. 3.7: DC responses of the complementary photoreceptors. V_{bias2} and V_{bias_ptype} bias voltages do not affect the slopes of the output voltages, but only a dc offsets.

A is the gain of the common source amplifier. In subthreshold region parasitic capacitances depend mostly on the size of the transistor as shown in Equation 3.4:

$$C_{gs} = WC_{ov} \tag{3.4}$$

This is because the conduction between drain and source comes from a diffusion, not from the channel formation, as it is in strong inversion. In order to limit the C_{gs} capacitance the transistors M1 in Figure 3.4 and in Figure 3.5 was chosen to be narrow. The second parasitic capacitance affecting bandwidth of the photoreceptor is the gate-drain capacitance of transistor M2. However, thanks to the cascade transistor M3, this capacitance is successfully limited. Transistor M4 acts as a current source. Based on the simulations, the choice of a such size of this transistor ensures that the common source amplifier will be biased properly. Additionally, V_{bias} gives freedom of adjusting a DC offset of the Photoreceptor. V_{cas} should be chosen such that the cascode transistor stays in saturation for a typical operation of the photoreceptors. Figures 3.9a and 3.9b show a frequency responses of the photoreceptors for different levels of photocurrents. Combining Equation 2.4 with Equation 2.19 the time constant τ_n becomes:

$$\tau_n = \frac{C_m}{\frac{I_{DS}}{nU_T}} = \frac{C_m n U_T}{I_{DS}} \tag{3.5}$$

The above equation indicates that the bandwidth is proportional to the photocurrents, what is illustrated by the frequency response of the photoreceptors in Figure 3.9a and Figure 3.9b . The arrow shows that with the decrease of the photocurrent, the bandwidth also decreases. In order to show clearly this dependency the responses were normalised, such that the DC gain is equal 1. Figures 3.9c and 3.9d show the same frequency response but now without a normalisation. Both P-type and N-type photoreceptors have almost identical frequency responses, what proves rational assumptions for the gain compensation. The gain depends

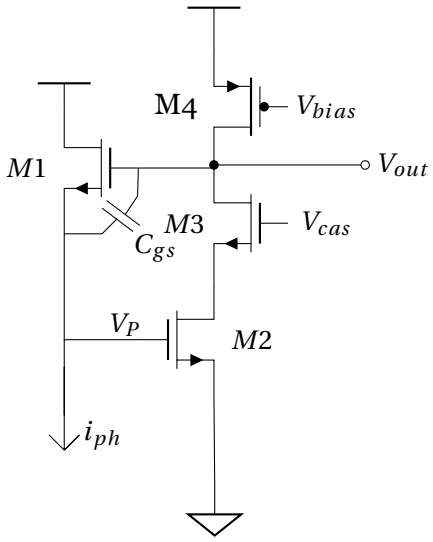


Fig. 3.8: Most significant parasitic capacitor affecting the bandwidth by the Miller Effect.

inversely on the photocurrents, such that highest gain is achieved when the transistor M1 is in deep subthreshold region. The subthreshold slope factor is also a function of the gate voltage[28]:

$$n = \left(1 - \frac{\gamma}{2\sqrt{V_G - V_{T0} + ((\frac{\gamma}{2}) + \sqrt{\Phi})^2}}\right)^{-1} \Rightarrow V_g \nearrow \longrightarrow n \searrow \tag{3.6}$$

For the higher photocurrent the feedback rises the gate voltage V_G of the transistor M1, what decreases the subthreshold slope factor.
Figure 3.10 shows a positive effect of a negative feedback. The bandwidth of the Photoreceptor implemented in the 3C-Pixel has over 4 decades higher bandwidth than the simple source follower photoreceptor from Figure 2.5. Only two photocurrent values are shown because for bigger photocurrents the source follower photoreceptor did not operate properly (the sourced transistor did not operate in subthreshold region any more).

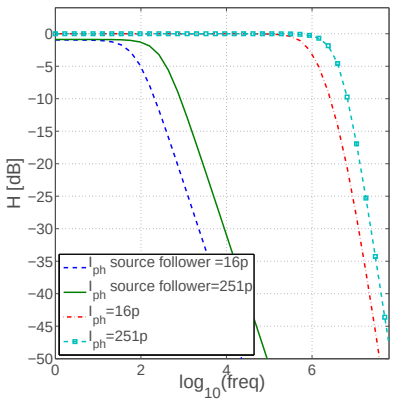
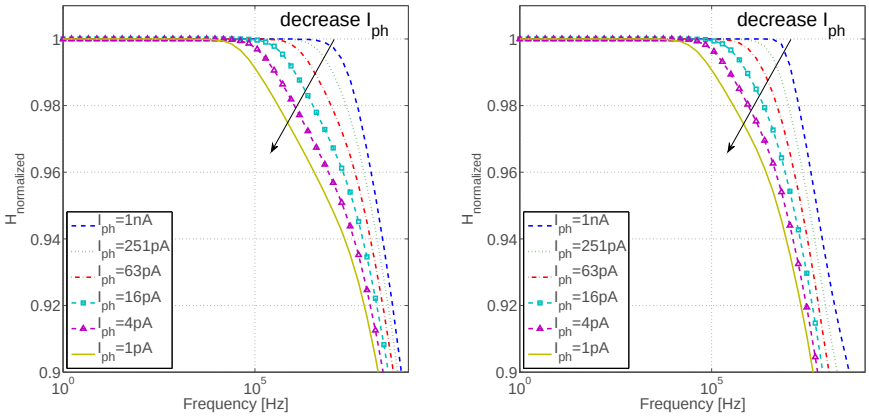


Fig. 3.10: Bandwidth comparison of the simple source follower and implemented photoreceptor.

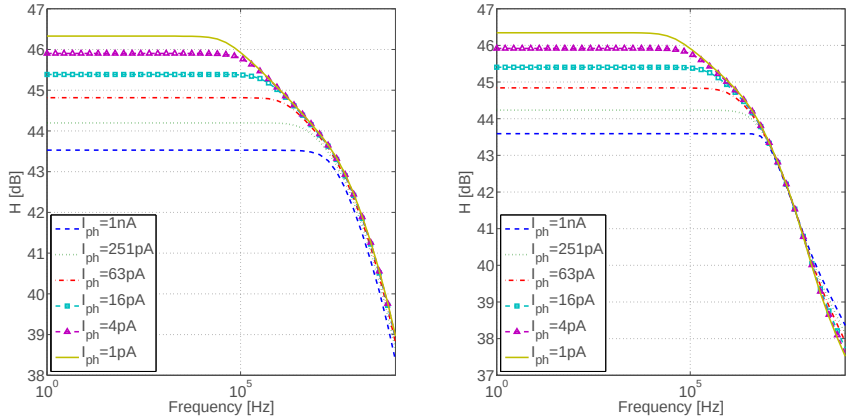
Proper biasing of the photodiodes

The three stacked photodiodes need a proper biasing to ensure that during the pixel operation they will remain reverse biased. The voltages: V_b, V_r and V_g from Figure 3.1 need to fulfill



(a) Normalised frequency response of the N-type Photoreceptor

(b) Normalised frequency response of the P-type Photoreceptor



(c) Frequency response of the N-type Photoreceptor

(d) Frequency response of the P-type Photoreceptor

Fig. 3.9: Frequency responses of the Photoreceptors.

the following conditions

$$V_b > V_g + 300mV \tag{3.7}$$

$$V_r > V_g + 300mV \tag{3.8}$$

From the DC analysis (Figure 3.24a) we know that the bias voltages V_{bias} and $V_{\text{bias\_ptype}}$ can be used to control DC levels in the photoreceptors. However this gives very limited freedom of control, because for too high bias voltages transistor M1 acting as a current source can leave a saturation region resulting in a non-linear gain. Another way of adjusting these voltages is to increase a source voltage of the transistor M2 in an N-type Photoreceptor and decrease in the P-type Photoreceptor. Any offset on the source side of the transistor M2 is reflected on the photoreceptor input. The simulations for different offsets suggested to set the source of the transistor M2 to 1.9V and 0.8V for the P-type and the N-type Photoreceptor respectively. This ensures that the stacked diodes being reverse biased and operating as a reliable photodiodes. The resulting voltages are now $V_b \approx 1.5V$, $V_g \approx 1.1V$ and $V_r \approx 1.5V$ Further, in Monte Carlo simulations the reverse bias voltages will be tested.

Stability

Figure 3.11 shows a phase margin as a function of bias voltage: V_{bias} and a background photocurrent I_{PH}. From Equation 2.14 we know that one of the dominating time constants depends on the transconductance of the transistor M2 from Figure 2.12. When V_{bias} on the PMOS transistor M4 increases, the current flowing through the transistor M2 decreases resulting in bigger time constant and bigger phase margin. For very high V_{bias} the phase margin drops below 20 degrees because the transistor M4 is almost not conducting, resulting in very big time constant. Phase

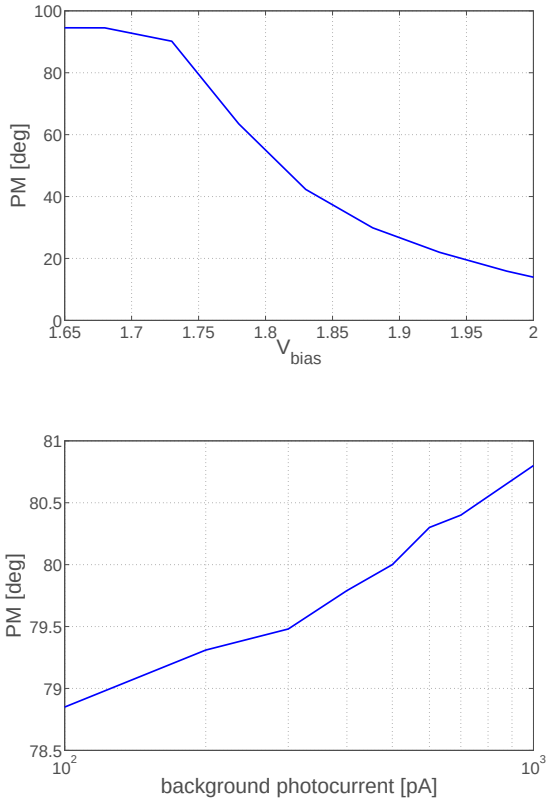


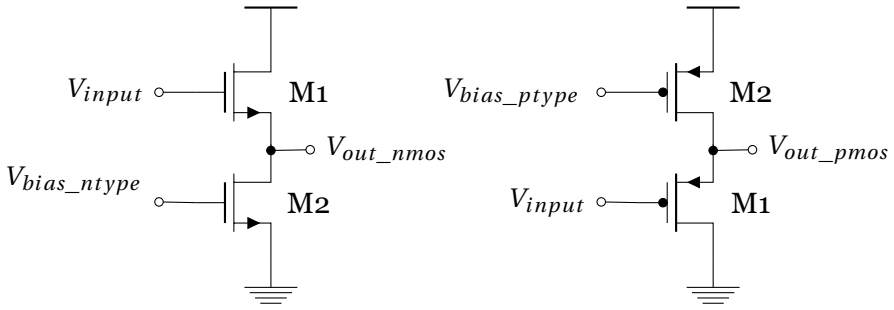
Fig. 3.11: Phase margin, as a function of a bias voltage: V_{bias} and a background photocurrent I_{PH}

margin of the photoreceptor depends also on the photocurrent, because as it was shown for the logarithmic photoreceptor in Equation 2.10, the time constant raising from the transistor in subthreshold region is inversely proportional to the background photocurrent. The dependency between phase margin and the background photocurrent is linear on the logarithmic X-axis because of the exponential dependency of the current and the gate-source voltage for transistor in subthreshold region.

3.1.4 Source Followers

During a reset mode the second stage amplifier produces high frequency noise affecting the output of the photoreceptor from Figure 3.4. The output of the photoreceptor is fed back through the transistor M1 to the input. As a

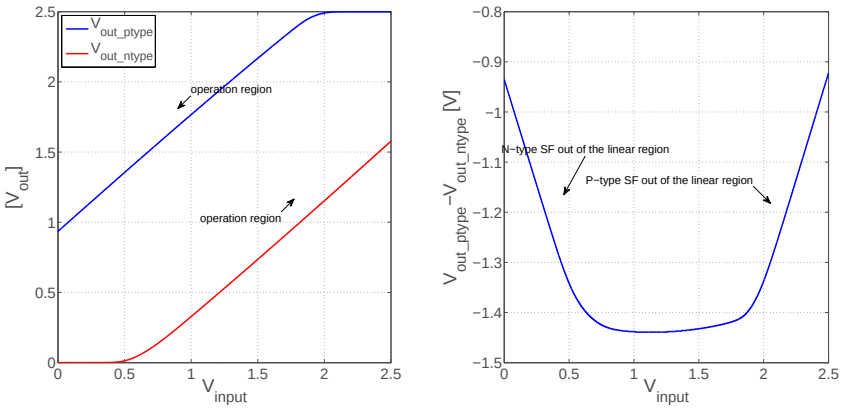
result, noise on the output of the Photoreceptor is reflected on its sensitive input, affecting the linearity of its gain. In order to limit this effect, the source followers (SF) from Figure 3.12 are used to isolate the differencing amplifier from the photoreceptor.



(a) N-type SF (b) P-type SF

Fig. 3.12: Schematic of the source follower. The bias voltages $V_{bias\_ptype} = 0.7V$ and $V_{bias\_ntype} = 1.9V$ keep the transistors M2 in saturation during the normal operation.

Linearity



(a) Comparison of the operation regions for P-type and N-type SFs

(b) Sum of V_{out_pmos} and V_{out_nmos}.

Fig. 3.13: Comparison of P-type and N-type SFs.

After adjustments aimed to obtain proper biasing of the photodiodes, the outputs of the N-type Photoreceptor (V_{o1} and V_{o3}) has raised to 1.6V - 2.2 V and the output of the P-type Photoreceptor (V_{o2}) has fallen to 0.7V - 1.1V. Since V_{o1} and V_{o3} output voltages are close to VDD it is better to use an N-type source follower to buffer them, whereas V_{o2} output should be buffered by a P-type source follower. The source followers do not operate properly until the input voltage is high enough (low enough for P-type SF) to supply (sink for P-type SF) the current equal to the current of the load transistor M2. Figure 3.13a illustrates the operation of P-type and N-type SFs. After this condition is met the source followers operate properly, what is indicated by the arrows. Since the source of the transistor M1 is not

connected to ground (and VDD for P-type SF), the Body Effect will make the threshold voltage to vary with the input voltage, resulting in lower linearity. The Body Effect for the P-type SF could be eliminated by connecting the NWELL of the transistor M1 to the output. However this is not a good solution, because it is not possible to connect a substrate to source of NMOS in N-type SF[2]. Hence, rational solution is to let the Body Effect affecting both types of SFs in the similar way such that the non-linearities of N-type and P-type SFs are similar. Finally, the most important is that all SFs attenuate the buffered signals V_{o1}, V_{o2} and V_{o3} in the same way.

3.1.5 Differencing amplifier

The second stage amplification is realised by an inverting amplifier with a feedback determined by a capacitive divider C_1/C_2. It only amplifies changes of the ratio of V_{o1}/V_{o2} and V_{o2}/V_{o3} [3]. As a result outputs V_{diffx} in Figure 3.1 indicate only if the change of the light illumination is *redder*, *bluer* whether *greener*.

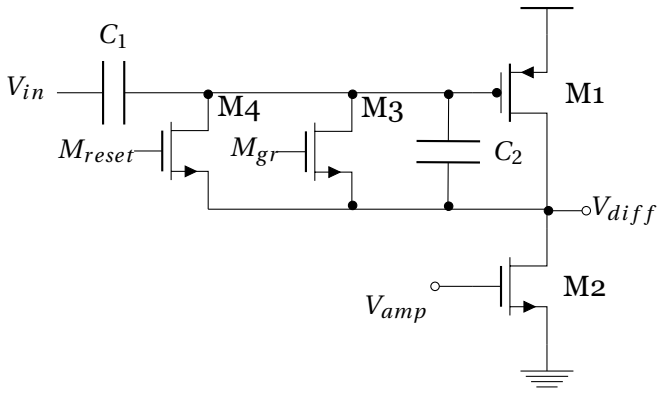


Fig. 3.14: A differencing amplifier.

Gain

The operation of the differencing amplifier can be explained by the abstracted schematic in Figure 3.15. Assuming infinitely high input impedance of the amplifier, the input current flowing through the capacitor C_1 will also flow through the capacitor C_2:

$$C_1 \frac{dV_{in}}{dt} = -C_2 \frac{dV_{diff}}{dt} \tag{3.9}$$

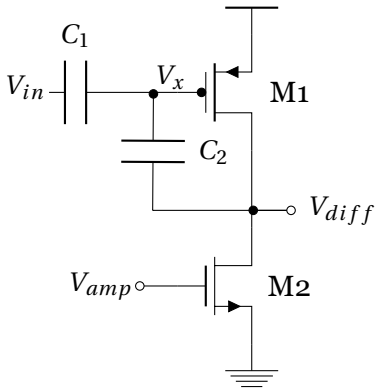
Equation 3.9 can be simplified:

$$C_1 dV_{in} = -C_2 dV_{diff} \tag{3.10}$$

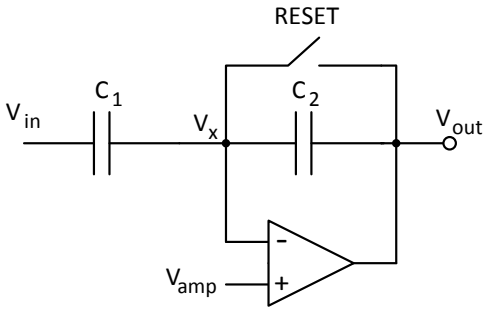
$$dV_{diff} = -\frac{C_1}{C_2} dV_{in} \tag{3.11}$$

[2]Deep-nwell could be used to separate the source of NMOS M1, but this costs additional space.

[3]Ratio V_{o1}/V_{o2} corresponds to the ratio between photocurrents $(I_b)/(I_b + I_g)$ and V_{o2}/V_{o3} to $(I_b + I_g)/(I_g + I_r)$



(a) Schematic.



(b) Abstracted schematic of the differencing amplifier

Fig. 3.15: Differencing amplifier normal operation.

Equation 3.11 indicated that the gain is determined by a ratio between capacitors: C_1/C_2. Since a DC voltage is cancelled by an input capacitor C_1, the differencing amplifier amplifies only changes of the input dV_{in}. It helps to achieve low mismatch between pixels because the DC mismatch coming from the first stage photoreceptors is removed.

Reset mode

When the differencing amplifier's output reaches a threshold value (ON or OFF), it is balanced by the RESET switch by shorting the input with the output. The amplifier should be balanced to the mid point between power rails, e.g. 1.25V in order to a provide symmetrical headroom for ON and OFF thresholds. In order to determine the sizes of the transistors M1 and M2, the operation of the differencing amplifier during RESET mode will be analysed. Based on Figure 3.16 current flowing through the transistors M1

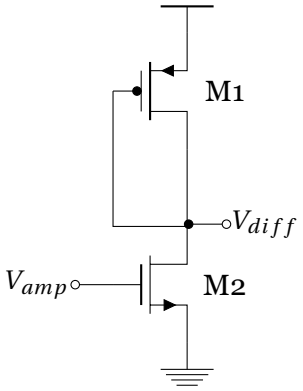


Fig. 3.16: Reset mode of the differencing amplifier.

and M2 are[4]:

$$\frac{1}{2}\frac{W_1}{L_1}\mu_p C_{ox}(|V_{GS_1}| - V_{Tp})^2 = \frac{1}{2}\frac{W_2}{L_2}\mu_n C_{ox}(|V_{GS_2}| - V_{Tn})^2 \qquad (3.12)$$

[4]Second order effects were omitted.

Further, the Equation can be simplified:

$$\frac{W_1}{L_1}\mu_p(|V_{GS_1}| - V_{Tp})^2 = \frac{W_2}{L_2}\mu_n(V_{GS_2} - V_{Tn})^2 \tag{3.13}$$

$$\frac{W_1}{L_1}\mu_p(|V_{diff} - VDD| - V_{Tp})^2 = \frac{W_2}{L_2}\mu_n(V_{amp} - V_{Tn})^2 \tag{3.14}$$

The threshold voltages from the Spice Model are extracted:

$$V_{Tn} = 0.51, V_{Tp} = 0.61 \tag{3.15}$$

The amplifier during reset should be balanced to 1.25V and V_{diff} for the convenience is chosen to be also 1.25V. Hence:

$$\frac{W_1}{L_1}\mu_p(1.25 - 0.61)^2 = \frac{W_2}{L_2}\mu_n(1.25 - 0.51)^2 \tag{3.16}$$

After rearranging Equation 3.17 becomes:

$$\frac{\frac{W_1}{L_1}}{\frac{W_2}{L_2}} = \frac{\mu_n}{\mu_p} \cdot 1.336 \tag{3.17}$$

In order to compare charge mobilities u_n and u_p, a transconductance parameter β was extracted for identically sized PMOS and NMOS transistors. Transconductance parameter β is:

$$\beta_{n/p} = \frac{1}{2}\frac{W}{L}\mu_{n/p}C_{ox} \tag{3.18}$$

For the same sizes of PMOS and NMOS transistors, the information about the ratio u_n/u_p is obtained in the following way:

$$\frac{\beta_n}{\beta_p} = \frac{\frac{1}{2}\frac{W_1}{L_1}\mu_n C_{ox}}{\frac{1}{2}\frac{W_2}{L_2}\mu_p C_{ox}} = \frac{u_n}{u_p} \tag{3.19}$$

From the simulation the ratio is $\frac{u_n}{u_p} = 3.267$. Including this results in the Equation 3.17:

$$\frac{\frac{W_1}{L_1}}{\frac{W_2}{L_2}} = 4.368 \tag{3.20}$$

Finally, taking into consideration the threshold difference and the charge mobility,the output voltage after reset will be equal 1.25V for the ratio $\frac{W_1/L_1}{W_2/L_2} = 4.368$. The sizes of the transistors were chosen according to these calculations and collected in the Table 3.4. The simulations confirmed sizes choice.

Transistor	W [um]	L [um]
M1	1.1	0.6
M2	1	2.3

Table 3.4: Sizes of the transistors of the Differencing Amplifier from Figure 3.14

Capacitor	Capacitance [fF]
C_1	416.432
C_2	70.98

Table 3.5: Specification of capacitors used in the design of a differencing amplifier.

Capacitors ratio

Capacitors values chosen in this design are shown in Table 3.5. Based on the Equation 3.11, differencing amplifier amplifies the changes of the input with a gain determined by a ratio C_1/C_2. It is desirable to create capacitor C_1 as big as possible and capacitor C_2 as small as possible in order to obtain high gain of the differencing amplifier. However if the capacitor C_2 is very small the mismatch becomes to big and the resulting gain of the pixel has to high deviation. The choice of $70.98fF$ appears to be optimal trade-off between the gain, the capacitors ratio mismatch and area. The capacitors were organised in common centroid structure described in *Layout* section.

3.1.6 Comparators

The differencing amplifier's output is compared against two threshold values (ON and OFF) by comparators. The comparator is realised by a common source amplifier shown in Figure 3.17. The gate of transistor M1 acts as an input, while the gate of transistor M2 is used to set the threshold value for the comparator. The sizes of transistors realising the common

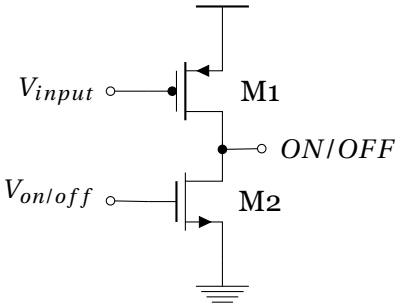


Fig. 3.17: Comparator - ON/OFF event generator.

source amplifier (Table 3.6) were chosen to be identical to the transistors realising a differencing amplifier from Figure 3.14. Recalling that in

Transistor	W [um]	L [um]
M1	1.100	0.6
M2	1	2.3

Table 3.6: Sizes of the transistors of comparator from Figure 3.17

order to provide a symmetrical headroom for ON and OFF thresholds, the differencing amplifier is balanced to 1.25V for V_{amp} = 1.25. Since the

comparator and the differencing amplifier are sized identically they have the same operating points, e.g. any increase or decrease of $V_{on/off} \pm V_{offset}$ will decrease weather increase the threshold of the comparator respectively. This is illustrated in Figure 3.18. Since the common source amplifier

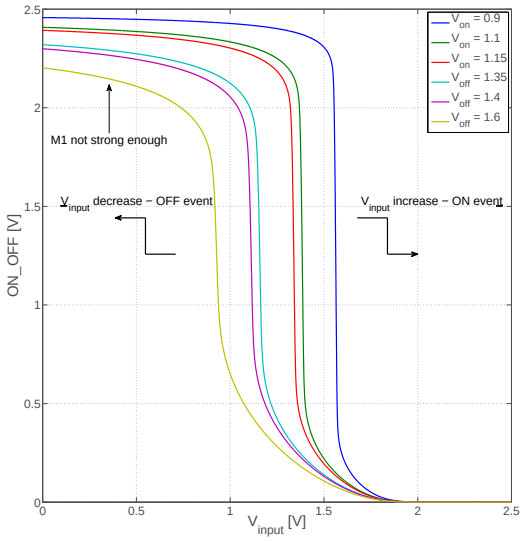


Fig. 3.18: ON/OFF vs. V_{input} for different $V_{on/off}$.

has high gain, the transition when one of the transistors becomes more conductive is very steep. The threshold value for which this transition occurs is shifted by $V_{on/off}$. For $V_{on/off} > 1.25V$ the threshold decreases, what is indicated by the direction of the arrow of OFF event. Similarly, for ON event the condition is $V_{on/off} < 1.25V$. As a result, ON event is indicated by the comparator by $VDD \rightarrow GND$ and OFF event by $GND \rightarrow VDD$ transitions. For high thresholds transistor M1 is not strong enough to pull up the output to 2.5V. However it does not cause any problems because OFF event is later pulled up by an inverter.

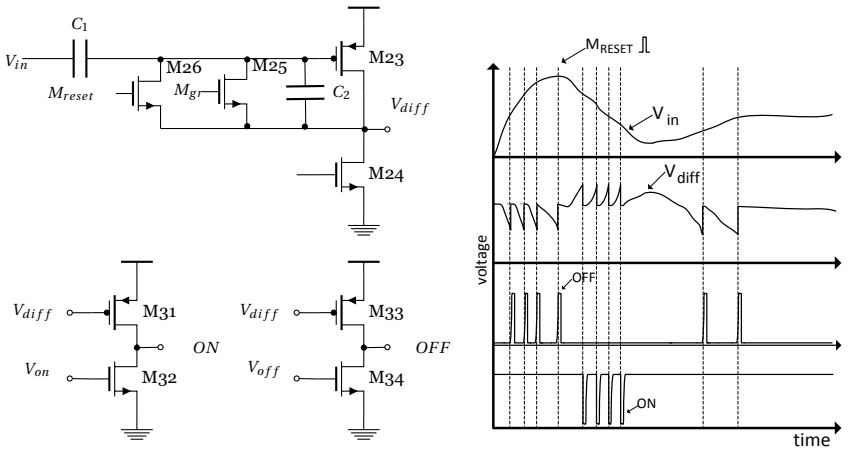
Differencing amplifier with comparators

Figure 3.19 explains the principle of operation of the differencing amplifier together with comparators. The entire process of handshake communication and arbitration in AER protocol is omitted in order to show a basic principle how the events are generated in 3C-Pixel. Timing diagram illustrates, that the steeper transition of the input signal is, the more events are generated. This is a critical stage for the 3C-Pixel performance, because for substantial mismatch in capacitor ratio or comparators' threshold the pixel will suffer from poor FPN[5]. The robustness against mismatch will be tested by *Monte Carlo* simulations in *Simulations* section.

3.1.7 3C-Pixel circuitry

So far we have described single building blocks of the pixel. Now the operation of the entire pixel will be described, with division into analog

[5]FPN in this design is described as a different number of ON/OFF events for the same stimuli



(a) Schematic.

(b) Timing diagram.

Fig. 3.19: Differencing amplifier operating with comparators. Any time V_{diff} exceeds a threshold, the comparator indicates it by an ON or OFF signal. A reset signal M_{reset} balances the differencing amplifier to 1.25V. Since the differencing amplifier is inverting, actually OFF event indicates increase and ON event decrease of the input V_{in}.

and digital parts. In the analog part changes between photo currents from the stacked photo diodes are amplified and compared against thresholds. The digital part handles the communication between the pixel and the AER peripheries described in chapter *Background*.

Analog part

Figure 3.20 shows analog circuitry of the 3C-Pixel. It consists of the blocks described in the previous sections: photodiodes, photoreceptors, source followers, differencing amplifiers and comparators. The front-end stage is realised by stacked photo diodes supplying three photocurrents: i_b, i_{b+g} and i_{r+g}. The changes of the photocurrents are transduced logarithmically by photoreceptors into V_{o1}, V_{o2} and V_{o3}:

$$\Delta V_{o1} = \Delta U_T \cdot n_1 ln(\frac{i_b}{I_{S1}}) \tag{3.21}$$

$$\Delta V_{o2} = -\Delta U_T \cdot n_8 ln(\frac{i_{b+g}}{I_{S8}}) \tag{3.22}$$

$$\Delta V_{o3} = \Delta U_T \cdot n_9 ln(\frac{i_{r+g}}{I_{S9}}) \tag{3.23}$$

I_{Sx} is the specific currents for the transistor in subthreshold region and n_x is a subthreshold slope factor. The negative sign for V_{o2} comes from the fact of using a P-type photoreceptor. These voltages are then buffered by the source followers resulting in:

$$\Delta V_{so1} = \Delta U_T \cdot \frac{n_1}{n_{13}} ln(\frac{i_b}{I_{S1}}) \tag{3.24}$$

$$\Delta V_{so2} = -\Delta U_T \cdot \frac{n_8}{n_{17}} ln(\frac{i_{b+g}}{I_{S8}}) \tag{3.25}$$

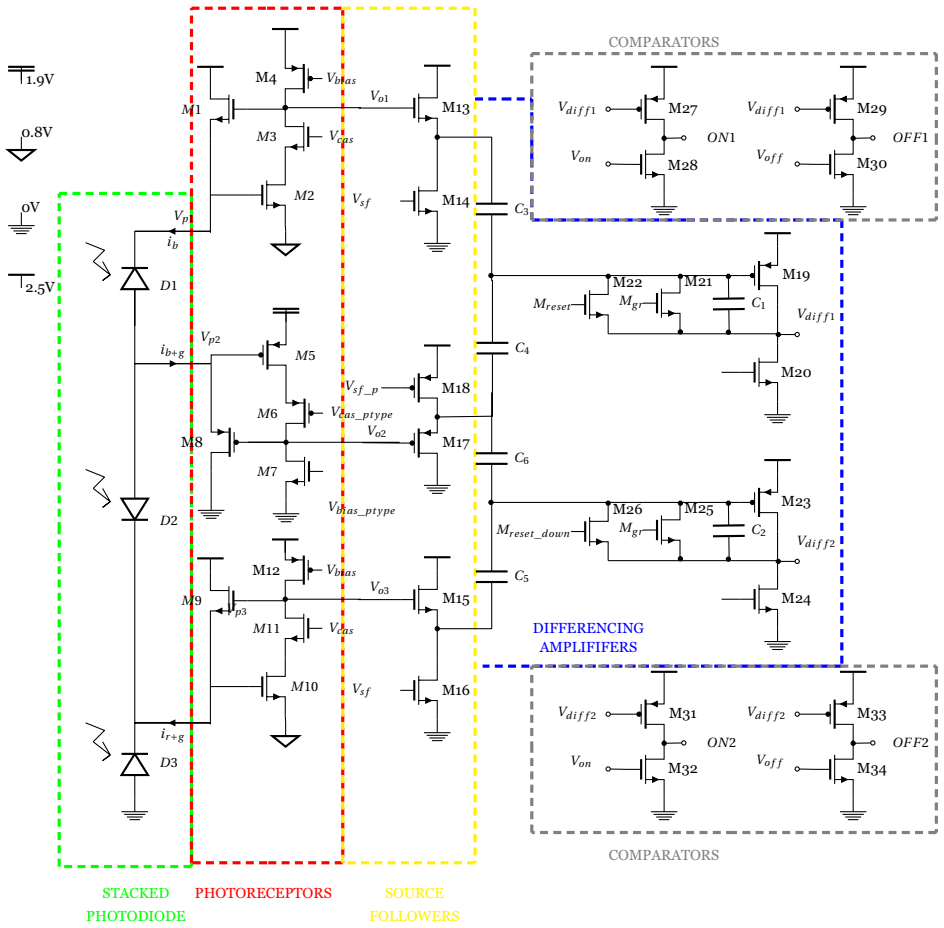


Fig. 3.20: Circuit of the 3C-pixel logarithmic photoreceptors amplifying three different photocurrents - analog part

$$\Delta V_{so3} = \Delta U_T \cdot \frac{n_9}{n_{15}} ln(\frac{i_{r+g}}{I_{S9}}) \tag{3.26}$$

The resulting pairs of changes of voltages $V_{so1} - V_{so2}$ and $V_{so2} - V_{so3}$ are amplified by the differencing amplifiers:

$$\Delta V_{diff1} = A_1 \Delta U_T \cdot \frac{n_1}{n_{13}} ln(\frac{i_b}{I_{S1}}) - A_2 \Delta U_T \cdot \frac{n_8}{n_{17}} ln(\frac{i_{b+g}}{I_{S8}}) \tag{3.27}$$

$$\Delta V_{diff2} = A_3 \Delta U_T \cdot \frac{n_9}{n_{15}} ln(\frac{i_{r+g}}{I_{S9}}) - A_4 \Delta U_T \cdot \frac{n_8}{n_{17}} ln(\frac{i_{b+g}}{I_{S8}}) \tag{3.28}$$

It can be proved that [10]:

$$A_1 \frac{n_1}{n_{13}} = A_2 \frac{n_8}{n_{17}} = A_3 \frac{n_9}{n_{15}} = A_4 \frac{n_8}{n_{17}} = A \tag{3.29}$$

This assumption is valid because the photoreceptors gains and source followers were compensated. Also the gain A_1, A_2, A_3 and A_4 is specified by a known and identical ratio between capacitors C_3/C_1, C_4/C_1, C_6/C_2 and

C_5/C_2 respectively. Equations 3.27 simplifies:

$$\Delta V_{diff1} = A \cdot U_T(\Delta ln(\frac{i_b}{I_{S1}}) - \Delta ln(\frac{i_{b+g}}{I_{S8}})) \\ = A \cdot U_T \cdot \Delta(ln(\frac{i_b}{i_{b+g}}) - ln(\frac{i_{S1}}{I_{S8}})) \\ = A \cdot U_T \cdot \Delta(ln(\frac{i_b}{i_{b+g}})) \quad (3.30)$$

Similarly ΔV_{diff2} becomes:

$$\Delta V_{diff2} = A \cdot U_T \cdot \Delta(ln(\frac{i_{r+g}}{i_{b+g}})) \quad (3.31)$$

Equations 3.30 and 3.31 indicate that the 3C-Pixel responds only to changes in the photocurrents ratios. A change in overall intensity without a change in color spectrum will keep the ratio i_{r+g}/i_{b+g} and i_b/i_{b+g} constant. The changes ΔV_{diff1} and ΔV_{diff2} are finally compared against thresholds adjusted by V_{on} and V_{off} in comparators, and if the threshold is exceeded ONx or OFFx event is generated.

Digital part

A digital circuitry is realised such that it can handle a communication with an AER periphery - Figure 3.21. Events ON1, ON2, OFF1 and OFF2 begin a cycle during which an address of the corresponding event is read out. The AER peripheries are explained in chapter *Background*. The bias voltage V_{refr} controls the refractory time of the pixel, by changing the conductivity of the pull-up transistor equivalent to the refractory time of the neuron. This is an additional time how long the pixel is in reset mode and it can not generate another request.

3.2 Color recognition in 3C-Pixel

Equations 3.30 and 3.31 indicate that the 3C-Pixel amplifies changes in ratio between i_b/i_{b+g} and i_{r+g}/i_{b+g} photocurrents. Table 3.7 shows how the changes in these two ratios are represented by ONx, OFFx events and how they can be translated into color changes. Even though the photocurrents are mixtures of different spectra we can simplify and assume those three spectra as red, green, and blue light and denote the filtered intensities as R, G, and B. Changes in spectra (R, G and B) and photo currents (I_{r+g}, I_{b+g} and I_b) can be now used interchangeably because of the linear dependence between these two quantities. We will check if these assumptions are valid in chapter *Experimental Results*, where the stacked photo diodes implemented in silicon are characterised in terms of spectral response.

3.3 Vision Sensor Design

The implemented vision sensor consisting of 16 raws and 16 columns of 3C-Pixels and AER peripherals. Further the vision sensor is called 3C-TVS (abbreviation from a Three Color Temporal Vision Sensor). A 3C-TVS periphery consists of an AER building blocks described in section

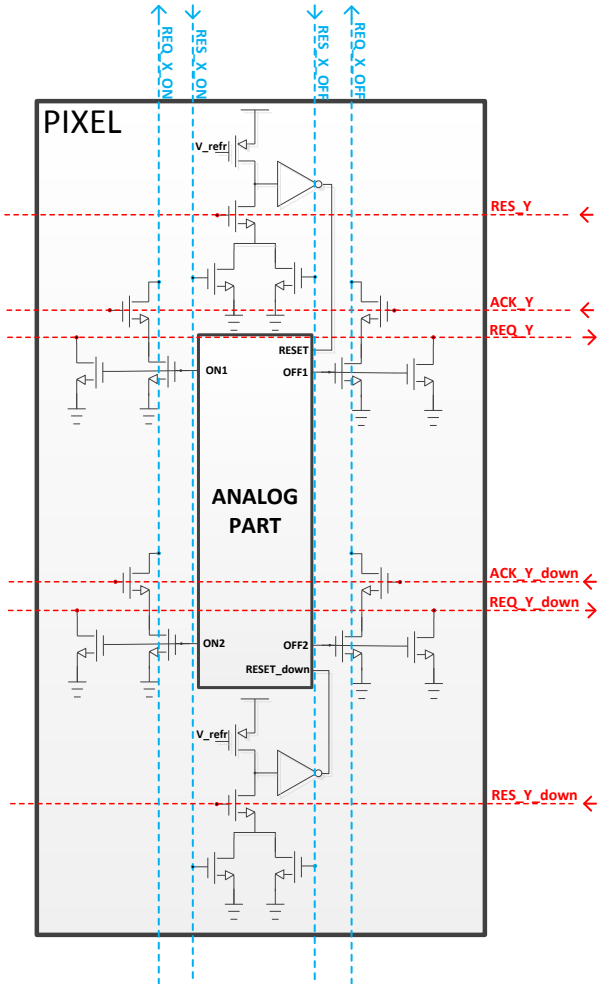


Fig. 3.21: Pixel handshaking internal circuitry. Pixel generates 4 outputs, which need a unique address: ON1, ON2, OFF1, OFF2.

Background. Since 4 different events are generated by a single 3C-Pixel, 1024 different addresses must be handled by the AER periphery.

3.3.1 Test 3C-Pixel and Test Stacked Photo Diodes

Additionally, a single test photo diodes structure and the separate test 3C-Pixel were implemented in order to characterise the stacked photo diodes spectral response and the behavior of a single pixel. The following signals are sensed in the test pixel: V_{o2}, V_{o3}, V_{diff1} and V_{diff2} (all analog signals) and ON/OFF events, M_{reset_down}, M_{reset} (digital signals). Analog signals are buffered by the unity gain amplifier and the digital signals are buffered by digital buffers. The unity gain amplifier is realised by the two stage amplifier from Figure 3.22, where V_{in-} is shorted with V_{out}. The designed amplifier was tested in terms of stability and current consumption, and it is summarised in Table 3.8. The digital buffers are realised by strong inverters. The digital circuitry of the test pixel is modified in order to omit a handshake communication with the AER peripheries. Pull down transistors, from Figure 3.21, generating REQ_Y and REQ_X are replaced with a simple 3 input *AND gate* from Figure 3.23. Thanks to this solution the pixel is self-reseting: it is being reset immediately after it sends an ON

Event	Photocurrents ratio	Assumed spectra ratio	Color change
ON1	$i_b/i_{b+g}\nearrow$	$\propto \frac{B}{G}\nearrow$	bluer
OFF1	$i_b/i_{b+g}\searrow$	$\propto \frac{B}{G}\searrow$	greener
ON2	$i_{r+g}/i_{b+g}\nearrow$	$\propto \frac{R}{G}\nearrow$	redder
OFF2	$i_{r+g}/i_{b+g}\searrow$	$\propto \frac{R}{G}\searrow$	greener

Table 3.7: Translation of ONx, OFFx events into color changes. The density of ONx/OFFx events is proportional to the rate of the color changes.

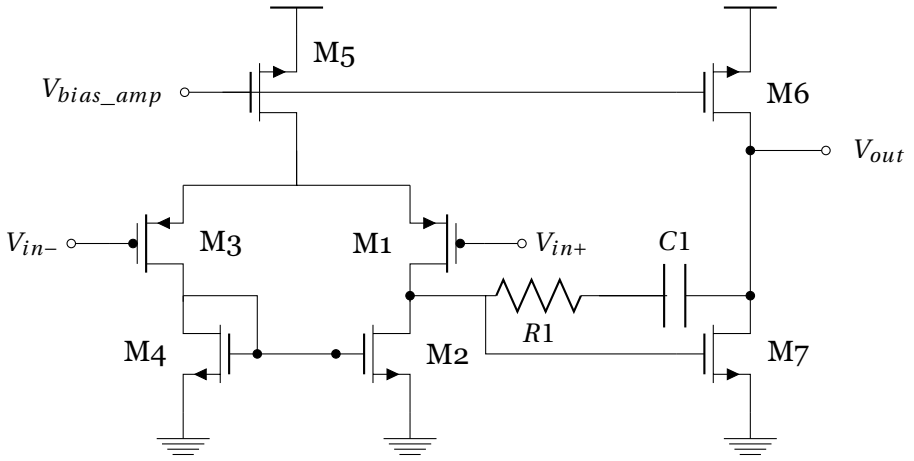


Fig. 3.22: Two stage amplifier with compensation, the design is based on the design guide from [37]

whether OFF event. The additional INIT input signal allows to keep the pixel in reset mode.

3.4 Bias voltages and interface pins

All bias voltages of the implemented vision sensor are connected to the external pads of the chip in order to be able to make off-chip bias adjustments. The bias voltages and power supplies with description and the values used during simulations are gathered in Table 3.9. The implemented chip has three separate power domains: for a pad frame with ESD protection, for buffers with a test pixel and the actual power for the vision sensor. Such solution enables to measure a power consumption of the designed vision sensor excluding irrelevant peripheries.

3.5 Simulations

3.5.1 3C-Pixel test

In order to test the 3C-Pixel response, a transient simulation was run with the photocurrents stimulus parameters included in Table 3.10. We stimulate the pixel only by a photocurrent I_r from a photo diode D3, what corresponds to illuminating the pixel surface by the ideal red light. Naturally, such conditions can not be achieved in real life because it is rare

Parameter	Value
Tail current	$10\mu A$
DC gain	90dB
Phase margin	52 deg
GB product	4MHz

Table 3.8: Specification of the two stage amplifier designed to realise a unity gain amplifier for buffering the analog outputs.

Fig. 3.23: Reset circuitry in a test 3C-Pixel.

to illuminate the pixel by so narrow band light. Secondly, the stacked photo diodes do not have such ideal response to the narrow band light and the experimental results will show later that the spectral responses of the photo diodes overlap. However, for test purposes such stimuli gives meaningful information about the symmetry of the response of the 3C-Pixel and gain compensations of the differencing amplifiers and source followers. From Equations 3.30 and 3.31 we know that outputs of differencing amplifiers are proportional to the ratio between photocurrents:

$$\Delta V_{diff1} =\propto \Delta \frac{i_b}{i_{b+g}} \tag{3.32}$$

Since there is no red photocurrent, the ratio in Equation 3.32 is always constant, meaning that there are expected no ON1/OFF1 events from the differencing amplifier no. 1. The second ratio is:

$$\Delta V_{diff2} =\propto \Delta \frac{i_{r+g}}{i_{b+g}} \tag{3.33}$$

For our stimuli, only a nominator is changing, hence the pixel should generate equal number of ON2 and OFF2 events.
Figure 3.24b shows outputs of three photoreceptors in the 3C-Pixel when stimulated with the stimuli from Figure 3.24b. Only V_{o3} follows the changes of I_r. Figure 3.25 shows the outputs of the differencing amplifiers: V_{diff1} and V_{diff2}, and digital outputs ON1, ON2, OFF1 and OFF2. As expected, only V_{diff2} generates events ON2 and OFF2. ON2 events indicate when the photocurrent has positive slope and OFF2 events indicate negative slope. Event though we kept the same thresholds for ON and OFF events (1.21V and 1.29V respectively) number of ON and OFF events is not the same (4 ON2 events and 6 OFF2 events for each cycle). The reason of non-symmetrical response is a charge injection from the reset transistor.

Charge injection

NMOS Transistors M22 and M26 in Figure 3.20 balance the differencing amplifiers by shorting their outputs with inputs. After balancing the gate

Bias voltage	Voltage [V]	Description
V_{ON}	< 1.25	threshold value for ON events
V_{OFF}	> 1.25	threshold value for OFF events
Refractory	≈ 1.9	refractory time control
V_{amp}	1.25	internal bias of the differencing amplifier
Tuned_vdd	1.9	P-type photoreceptor internal bias to ensure proper biasing of the photodiode
Tuned_gnd	0.8	N-type photoreceptor internal bias to ensure proper biasing of the photodiode
V_{sf}	0.7	Source follower N-type biasing
V_{sf_p}	1.9	Source follower P-type biasing
V_{bias}	1.7	N-type photoreceptor internal
V_{bias_ptype}	0.7	P-type photoreceptor internal
V_{cas}	1.6	N-type photoreceptor cascode transistor
V_{cas_ptype}	1	P-type photoreceptor cascode transistor
Pu_{bias_frame}	1.9	AER periphery pull-up transistors
Pd_{bias_frame}	0.6	AER periphery pull-down transistors
$VDD_{padframe}$	2.5	Power supply for a chip pad frame
$VDD_{Buffers}$	2.5	Power supply for analog and digital buffers
$VDD_{VisionSensor}$	2.5	Power supply for the vision sensor

Table 3.9: Bias voltages and power supplies used during simulations of the 3C-TVS Vision Sensor.

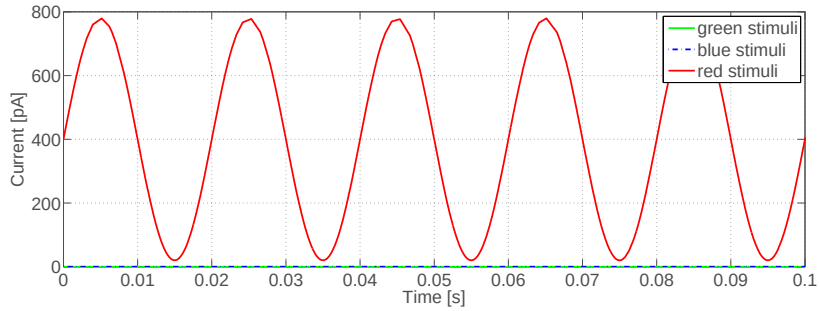
Diode	Photocurrent	DC [pA]	Amplitude [pA]	Frequency [Hz]
D1 - blue	I_b	0	0	-
D2 - green	I_g	0	0	-
D3 - red	I_r	400	380	50

Table 3.10: Photocurrents stimulus used to test the 3C-Pixel response.

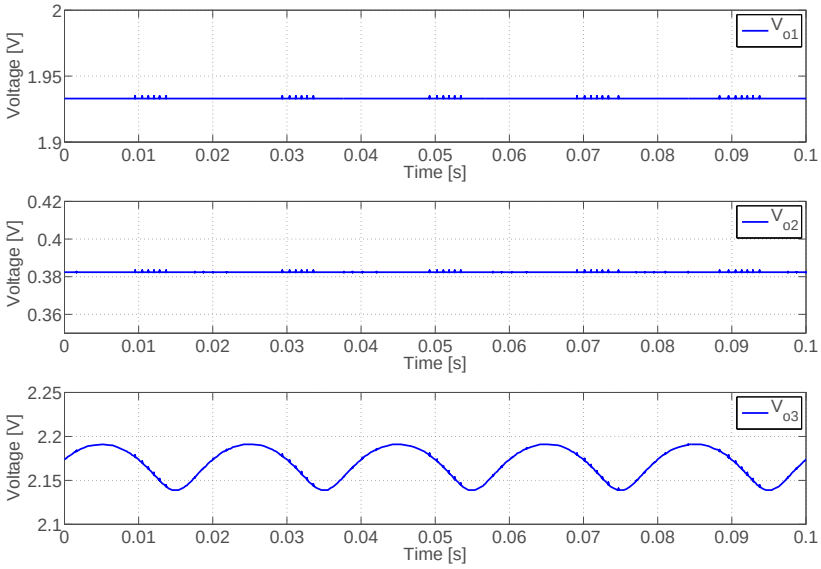
voltage of these transistors moves from $2.5V$ to $0V$ in order to turn them off. During switching unwanted negative charge is injected into the inverting input of the amplifier. As a result output voltages V_{diff1} and V_{diff2} move approximately $3mV$ after balancing, as it is shown in Figure 3.26. This results in statistical offset and higher sensitivity of the pixel to OFF events, shown later in Monte Carlo Simulations. This is not a substantial problem, because this offset can be compensated by proper adjusting thresholds.

Monte Carlo Simulations

While designing the 3C-Pixel, each block was simulated including Monte Carlo simulations. Especially, it was a case to compare responses to mismatch of complementary circuits, e.g. P-type and N-type photoreceptors and source followers. However, due to staggering amount of data results including Monte Carlo simulations of each single block are omitted in this thesis and most meaningful results including Monte Carlo simulations of entire pixel are presented.
The 3C-Pixel was stimulated with the same photocurrents stimulus as in the previous simulation (Table 3.10). Hence, again we expect only ON2 and OFF2 events. Threshold values V_{on} and V_{off} were set to 1.3V and 1.2V



(a) Stimuli



(b) Outputs of three photoreceptors in the 3C-Pixel when stimulated with the photocurrent corresponding to illumination by the ideal blue spectrum light.

Fig. 3.24: Response of the photoreceptors to one photocurrent stimuli.

respectively. A 400ms transient simulation was run in order to record 20 cycles of the sinusoidal stimuli used to characterise the performance of the pixel. Further, the analysis was repeated 1000 times including process mismatch. The results were normalized to one sinus period.

ON1+OFF1 vs. ON2+OFF2 events

Figure 3.27 shows histogram and Table 3.11 includes mean values and standard deviations of ON+OFF events from both differencing amplifiers[6]. As expected, staggering number of iterations gave zero ON1+OFF1 events, namely in 95% of MC iterations there were no events registered. Also, a low standard deviation indicates low spread of ON1+OFF1 events. Numbers of ON1+OFF1 events in histogram 3.28a are not integers, because of the normalisation of results to one sinus period. It means that the pixel needs few cycles to generate one event. The differencing amplifier no. 2 generates ON2+OFF2 events with a mean value of 10.45 and standard deviation of 0.4 event per sinus cycle. Again, a low standard deviation indicates low

[6] # - number of.

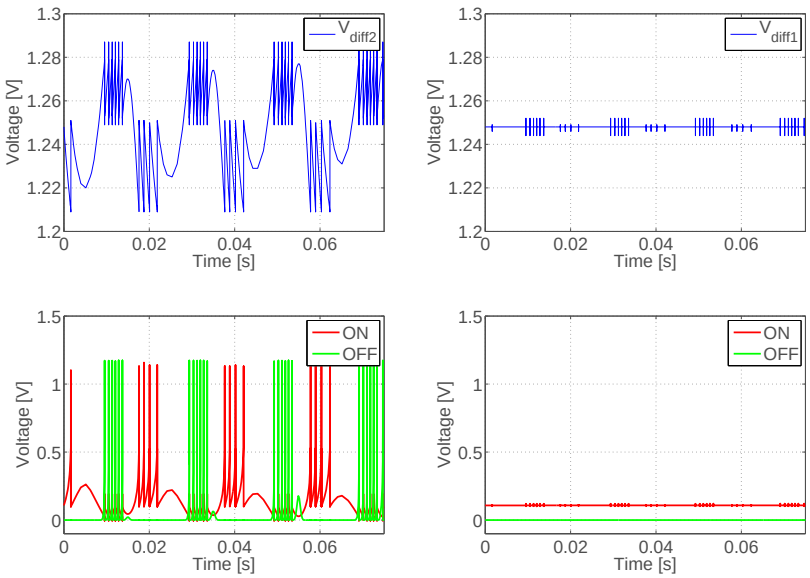


Fig. 3.25: The pixel response when stimulated with the photocurrent corresponding to illumination by the ideal blue spectrum light.

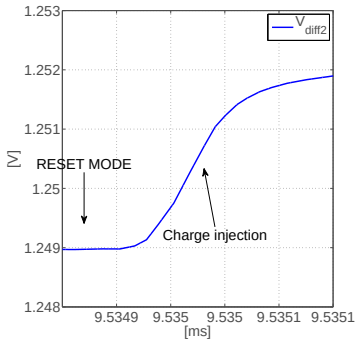


Fig. 3.26: Close look up on the output V_{diff2} after balancing reveals an effect of charge injection from the reset transistor M26.

spread of ON2+OFF2 events. The *ON1+OFF1 vs. ON2+OFF2 events* test proves the 3C-Pixel robustness against mismatch in terms of ability in detecting color intensity changes between red and blue color: i_g/i_b and discriminating a change in overall intensity: i_b/i_b. No spikes registered from the differencing amplifier no.1 prove correct gain compensation procedure of the photoreceptors and good design of the source followers.

ON vs. OFF events

Figure 3.28 shows histograms and Table 3.12 includes mean values and standard deviations of ONx and OFFx events from both differencing amplifiers. This test shows closer look on the symmetry of the response of the 3C-Pixel, e.g. equal number of ON2 and OFF2 events. As expected, a staggering number of iterations gave zero ON1/OFF1 events, namely 96.2% of ON1 events and 96.5% of OFF1 events were equal zero. Also, a low standard deviation indicates low spread of ON1 and OFF1 events. The

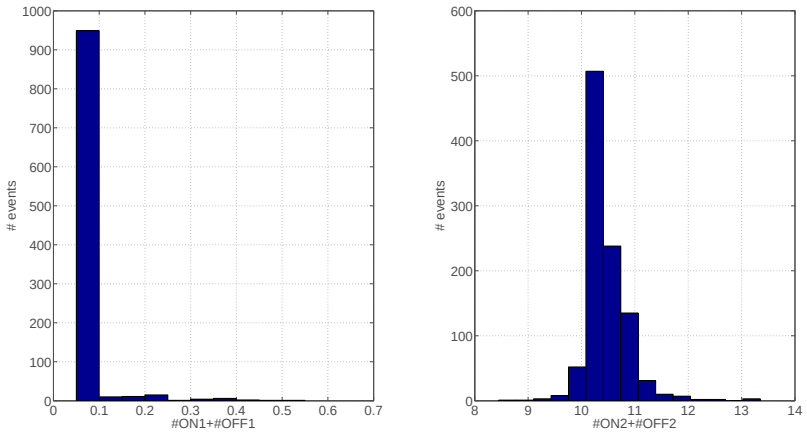


Fig. 3.27: Monte Carlo simulation, count of ON+OFF events. Number of iterations 1000.

	mean #	σ #
ON1+OFF1	0.081	0.054
ON2+OFF2	10.45	0.402

Table 3.11: Statistical properties of data shown in Histogram from Figure 3.27.

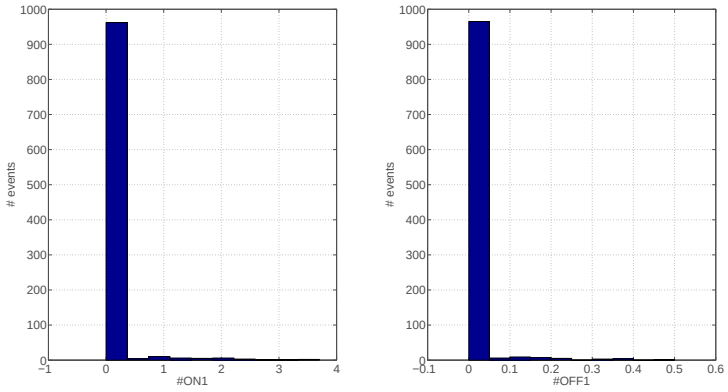
differencing amplifier no. 2 generates OFF2 events with a mean value of 5.54 and ON2 events with 5.1 events per sinus cycle with a low standard deviation.

	mean #	σ #
OFF1	0.008	0.045
ON1	0.050	0.310
OFF2	5.5315	0.2662
ON2	5.105	0.2413

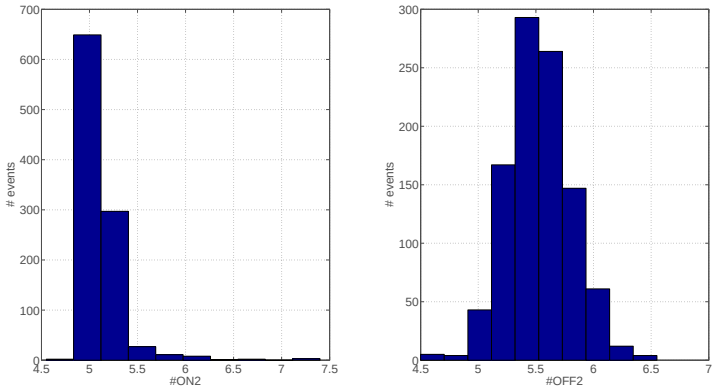
Table 3.12: Statistical properties of data shown in Histogram from Figure 3.28.

Histogram in Figure 3.29 illustrates the symmetry of the response of the 3C-Pixel in form of the ratio between OFF2 and ON2 events. Ideally, the ratio should be equal 1, meaning that number of ON2 and OFF2 events is equal. In fact, average ratio is 1.125, meaning that the pixel is more sensitive to the negative changes [7]. Not uniform distribution of ON and OFF events is explained by a charge injection from a reset transistor after a reset mode, as explained in previous section. The charge injected by a reset transistor after a reset mode increases the probability of OFFx event occurrence. The *ON vs.OFF events* test shows that the 3C-Pixel robustness against mismatch in terms of uniform distribution of positive and negative events. The statistical constant difference between number of ON2 and OFF2 events is very small and does not affect the performance of the pixel substantially. What is more, this problem can be easily solved by manipulating threshold

[7]Since the differencing amplifier is inverting OFF event means increase in ratio, while ON event means decrease in colour ratio.



(a) Number of ON/OFF events from the differencing amplifier no.1.



(b) Number of ON/OFF events from the differencing amplifier no.2.

Fig. 3.28: Monte Carlo simulation, count of ON/OFF events. Number of iterations 1000.

voltages.

Reverse biasing of the photodiodes.

Monte Carlo simulation was run in order to test if the three photodiodes in the stacked photodiode remain properly biased when the process mismatch is included. The resulting histogram is shown in Figure 3.30. The conditions assumed in Equations 3.7 and 3.8 are met, because the photodiodes reverse bias voltage does not drop below 300mV. The diode D1 has much higher reverse bias voltage, because its anode is connected to ground.

3.5.2 AER periphery test

In order to test the AER peripheries, a test bench consisting of 16 pixels and the AER peripheries was created. Figure 3.31 illustrates a test bench setup with position of stimulated pixels. The columns encode the event type: ON whether OFF and rows indicate the number of pair: 1 whether 2. Since both stimuli are an ideal red light, we expect from stimuli 1:
·000110 - ON2
·001110 - OFF2

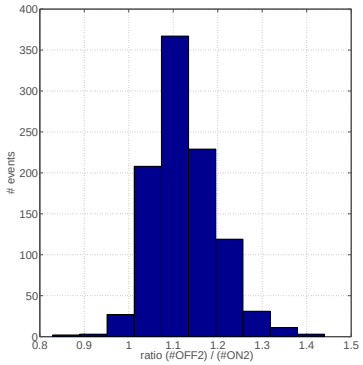


Fig. 3.29: Monte Carlo simulation, ratio between ON2/OFF2 events.

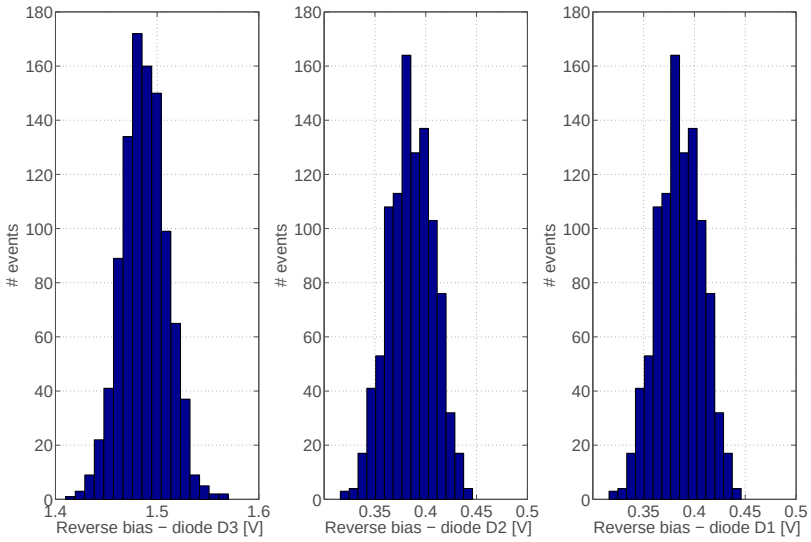


Fig. 3.30: Monte Carlo simulation, value of the reverse bias voltage in the stacked photodiode.

and from stimuli 2:
·100010 - ON2
·101010 - OFF2.

Figure 3.32 shows both stimuli and registered output bits. As expected, obtained addresses are 000110, 001110, 101010 and 100010.

3.6 Layout

3.6.1 Good layout guidelines

Even if the simulation results are very satisfactory, bad layout can worsen the results or, in the worst case, make the circuit stop working. Hence, the following general layout guidelines were taken into consideration while drawing layout:

1. Good methodology is to localise most sensitive and critical parts in the circuit, design them with considerable attention and isolate them from the noisy parts. Especially good approach is to separate the analog part from

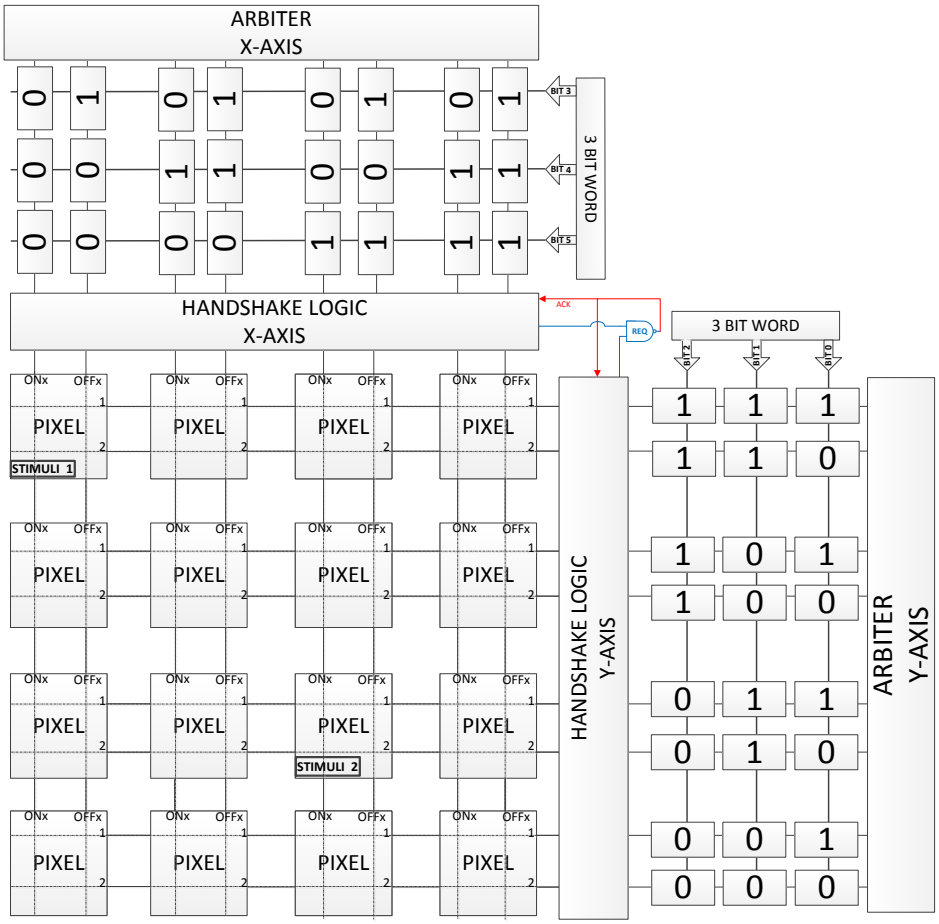


Fig. 3.31: Simplified test bench of the AER periphery with an array of 4x4 3C-Pixels. Stimuli 2 - only red photocurrent, Stimuli 1 - only red photocurrent

the noisy digital circuitry, and if possible, provide them separate ground pads to avoid substrate coupling between them. Another good method is to shield sensitive circuitry by guard rings, which later are properly terminated.

2. When designing an array of pixels, power supply distribution must be taken into consideration: if the metal wires distributing power in the array are not placed carefully, voltage drop can result in unwanted power gradient through the pixels, affecting their performance. Similar problems can cause not sufficient number of contact ties to substrate, resulting in not uniform ground potential reference for pixels. Also, tie contacts and wires widths have to be carefully considered to provide sufficient paths to sink currents. 3. The wires transmitting sensitive analog signals need to be properly isolated/guarded to avoid loss of information and digital signals need to be sufficiently buffered if they are transmitted for long distances. Good methodology is to minimise distribution of critical signals.

4. Wherever it is possible, decoupling capacitors should be placed in the layout. Decoupling capacitors help in reducing effect of the noise by high pass filtering it and, hence, reduce substrate coupling.

5. A common centroid structures should be used to ensure that the process/temperature gradients have zero net effect on device performance

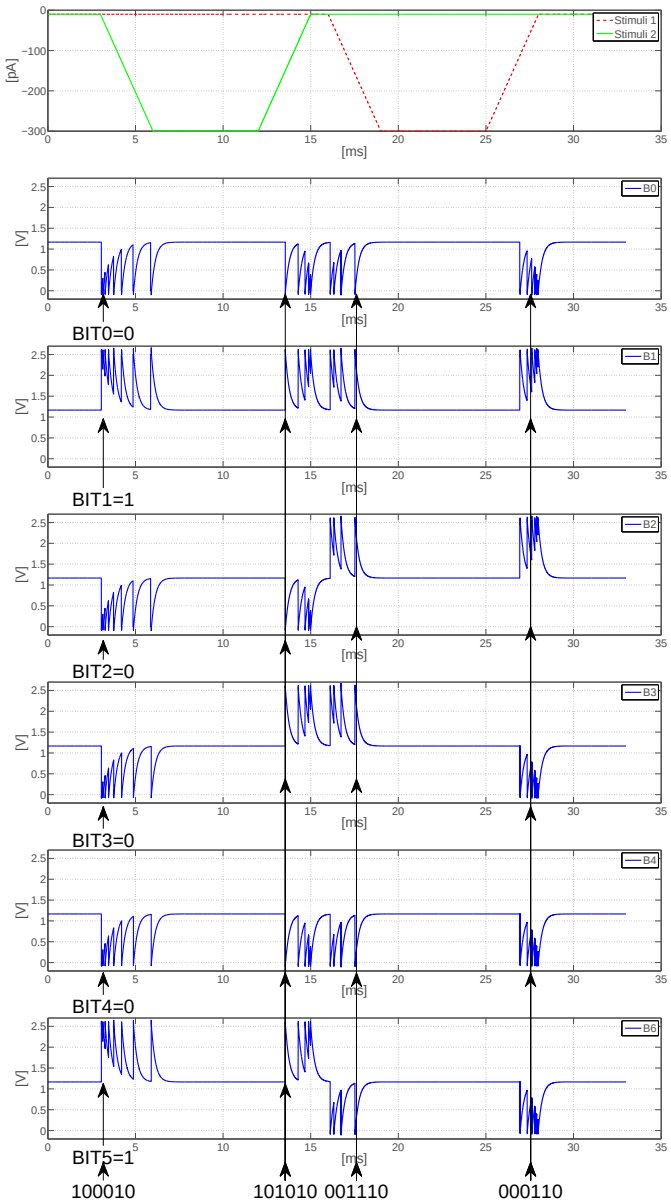


Fig. 3.32: Top plot:stimuli currents, the remaining plots: the corresponding addresses read out from the 4x4 test array.

[37]. This applies especially to the well-defined capacitance, resistance or transistor sizes ratios.

3.6.2 Stacked photodiode

Layout of the stacked photodiode is shown in Figure 3.33. Including a guard ring around, the size of the photodiode is $52\mu mx52\mu m$, with size $43.3\mu mx41.4\mu m$ of a photoactive area. In order to ensure that the photoactive area will not be covered by *dummies*, a special layer *dummy exclude* was used. During a typical CMOS process flow, one of the steps covers silicidation in order to lower the sheet resistance of drains and sources of transistors. This step was omitted for a photoactive area of the photodiode, because a silicide is highly opaque it would worsen a quantum efficiency of the implemented photodiode. The layout

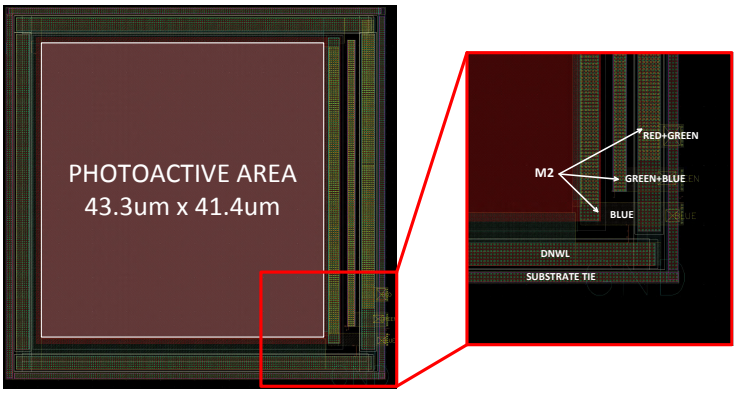


Fig. 3.33: Layout of the stacked photodiode. Metals M3-M9 are not shown for clarity.

is straightforward with the schematic of the stacked photodiode from Figure 2.4. Photocurrents from the three stacked photodiodes are routed on the right side in metal M2. The column closest to the photoactive area conducts current from $N+/PW$ junction, the second column from $N+/PW$ and PW/DNWL and the most right one from PW/DNWL and DNWL/PSUB.

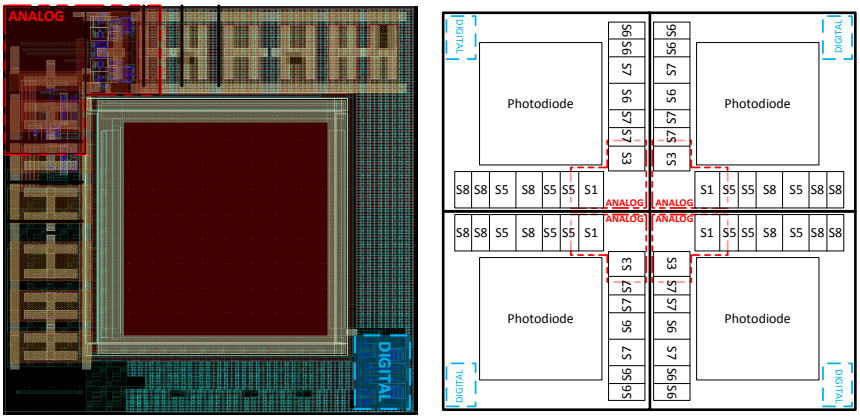
3.6.3 3C-Pixel

Figure 3.34b shows a the photoactive area occupation of one pixel, location of digital and analog circuitry and organisation of capacitors segments.

Capacitor*	**Segments****	**Capacitance** $[fF]$
C_1	S1	70.98
C_2	S3	70.98
C_3	S5	$161.42+2\cdot127.51$
C_4	S8	$161.42+2\cdot127.51$
C_5	S7	$161.42+2\cdot127.51$
C_6	S6	$161.42+2\cdot127.51$

Table 3.13: Segment organisation of capacitors used in the design of a differencing amplifier. $*$ - number of capacitors from Figure 3.6, $**$ - capacitors segments from Figure 3.34b

Pixel contains 6 capacitors (schematic Figure 3.6). These capacitors are divided into segments and organised in a centroid-like structure. Such organisation should help to make the capacitors ratios more robust against process mismatch and temperature gradients through the pixels array. Also, analog and digital circuitry of four adjacent pixels are localised in groups with a longest possible distance from each other in order to limit the effect of the digital noise on the sensitive analog signals. Table 3.13 shows MIM capacitors sizes and segments organisation.



(a) Single pixel layout. (b) Common centroid organisation.

Metal organisation

Thanks to a large number of metal layers (M1-M9) available in TSMC 90nm, different functions were assigned to different metal layers to properly organise signals and power distribution. At the end of this chapter a summary of each metal function in the matrix is provided (Table 3.16 and Table 3.17). Mainly, bias voltages and AER communication signals are routed by horizontal and vertical wires in different layers. In order to a provide uniform power distribution in the matrix, metals from each layer are used to distribute VDD. In order to meet a maximum density rules, a special square structure described in Tables 3.16 and 3.17 was designed. Such solution enabled a flexible creation of vast areas of metals meeting density rules.

3.6.4 Matrix and AER Peripheries

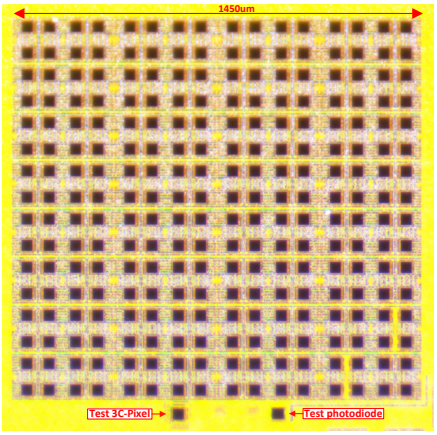
An array of 16x16 pixels organised in common centroid structure was created. Around the matrix peripheries responsible for handling the AER communication were implemented. Full chip layout is shown in Appendix C.

3.6.5 Microscope on-chip cell images

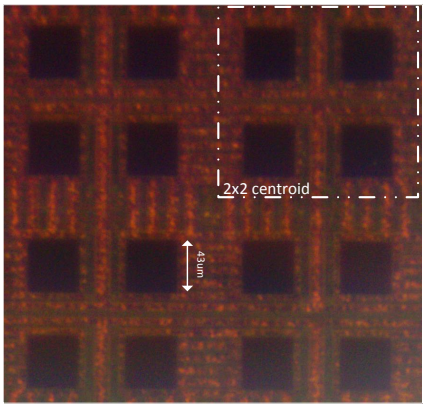
A microscope with zoom lenses and high resolution camera was used to make on-chip cell images. Figure 3.35a a chip surface. Locations of the test photodiode and the test pixel are indicated by arrows. Black squares represent photoactive area. There is no metal wiring nor dummy structures over the photodiodes in order to keep the quantum efficiency as high as possible. Figure 3.35b shows a closer look on pixels, we can distinguish 4 common centroids. Bronze horizontal and vertical lines across the array are top metal M9 lines distributing power.

3.7 Chip summary

Tables 3.15 and 3.14 summarise the chip size, pixels dimensions, analog and digital area occupation. The fill factor is reasonably high for such type of pixel (with processing circuitry inside the pixel). The fill factor is



(a) An array consisting of 16x16 3C-pixels, a single Test 3c-Pixel (bottom-left) and a single Test Photodiodes (bottom-right).



(b) 4 common centroid pixel structures. A a stacked photodiode size $43umx43um$. Bronze horizontal and vertical lines across the array are M9 lines distributing power.

limited because of the used MIM capacitors. Even though MIM capacitors

Chip size	$1394\mu x1394\mu$
Array size	$1312\mu x1338\mu$

Table 3.14: Chip dimensions, technology TSMC 90nm rf.

are located over the circuity in top metals and their sizes are the smallest possible which met the Monte Carlo simulations, they occupy almost 40% of the pixel area and limit further fill factor improvement. The alternative capacitor type which could be used is a MOM (Metal-Oxide-Metal) capacitor. In the used process MOM capacitors can create a multi-metal structures resulting in higher capacitance per area. Appendix D shows an alternative layout of the pixel with MOM capacitors used. However finally this design was not implemented in the array. The reason is the fact that MOM capacitors occupied six top metal layers and there was not enough place for proper wires routing. As a result wires transmitting digital and analog signals were to thin and the post-layout simulations were not satisfactory.

Pixel size	$82\mu x82\mu$
Photodiode size	$41.5\mu x43.4\mu$
Fill factor [%]	27
Area digital part [%]	1
Area analog part [%]	13
Area *MIM* capacitors [%]	39

Table 3.15: Pixel specification.

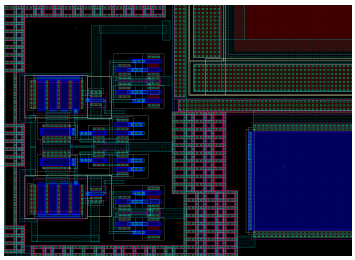
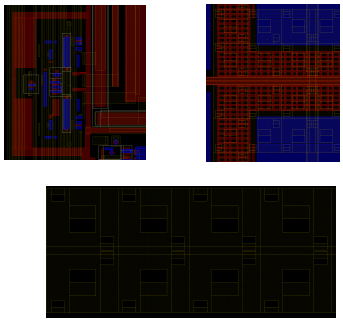
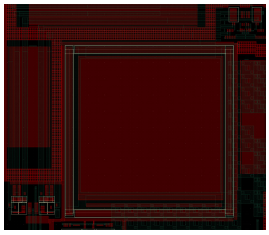
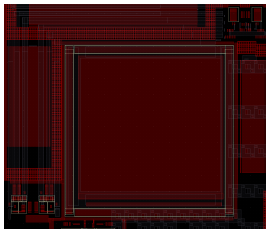
Mx	Usage	Example
M1	M1 is used for internal local connections in analog and digital circuitry of the pixel, as shown in the figure on the right side. The right side of the figure shows part of the decoupling MOS capacitor. The digital and analog circuitry is guarded by substrate contacts organised net structures connected by M1. Such solution always meets M1 maximum density rules.	
M2	M2 is used to distribute bias voltages along vertical direction in whole array, as it is shown in the first Figure on the right side. Also, M2 is used to distribute VDD in the array by the structure shown in the second and third Figure. A designed square in M2 has three rectangular holes to ensure to meet maximum density rules, when creating extensive structures distributing VDD in the array.	
M3	M3 is used to realise internal routing between distantly located digital and analog circuitry in the pixel. They conduct 4 signals: ON1, OFF1, ON2, OFF2. The width of the wires is approximately $1.8\mu m$. Also, identical square structure as in M2 is created to distribute power in the array.	
M4	M4 is also used to realise internal routing between digital and analog circuitry. Hence, signals ON1, OFF1, ON2, OFF2 are routed identically as in M3. M3 is also used to distribute ground in the array in order to create VDD-GND *a sandwich* structure for signals above and in order to create additional parasitic decoupling capacitance.	

Table 3.16: Metals M1-M4 usage. Figures show sections of the 3C-Pixel. For clarity only described metal is turned on in the Figures.

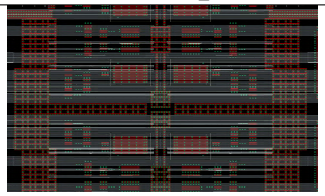
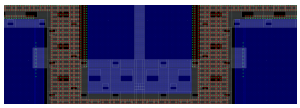
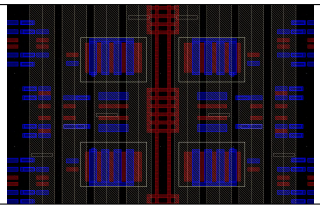
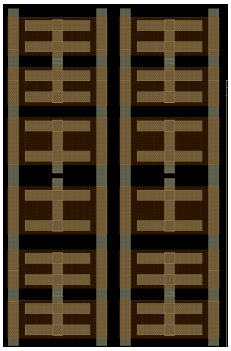
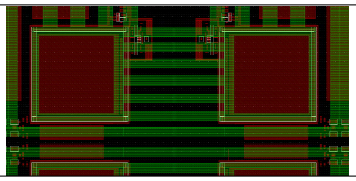
Mx	Usage	Example
M5	Metal M5 is used to route between pixels and AER peripheries in horizontal direction - long horizontal lines in the first Figure are REQ, ACK and RES signals handling handshake communication of two pixels in Y-axis. M5 is used also to distribute GND.	 
M6	Metal M6 is used to realise handshake communication in vertical direction and internal connections between *MIM* capacitors in the pixels.	
M7	Metal M7 is used for internal routing in the higher levels of the pixel, especially between *MIM* capacitors.	
M8	Metal M8 is used to create *MIM* capacitors together with Capacitor Bottom Metal layer and Capacitor Top Metal Layer. The Figure shows two set of of capacitors of the adjacent pixels in matrix. CBM layer is illustrated by orange colour in the Figure. M8 wires are wide, because minimum width rules are very restricted.	
M9	M9 realises power supply distribution. Similarly to M8, wires routed in M9 are wide, because minimum width rules are very restricted.	

Table 3.17: Metals M5-M9 usage. Figures show sections of the 3C-Pixel. For clarity only described metal is turned on in the Figures.

Chapter 4

Experimental Results

4.1 Laboratory Setup

The designed chip was fabricated and bonded in PLCC84 package. A custom PCB was designed (Appendix *Test PCB Layout*) in order to install a chip into a socket and carry out measurements. The test PCB contains:
· chip socket PLCC84,
· interface pins,
· reference bias voltages generators realised by potentiometers,
· test outputs realised by pinout and BNC connectors,
· separate power pins for 3C-TVS vision sensor, buffers and pad frame ESD protection,
· IDC-40 bus for AER communication with a data logger. The test PCB enables to setup all bias voltages. The fabricated vision sensor was charactesized using different laboratory setups.

4.1.1 3C-Pixel and photodiode measurement setup

In order to characterise a single test 3C-Pixel a setup able to dynamically change input colors was built. Two high-power full color LEDs (F50360) from Seoul Semiconductor were used described in Table 4.1. Two F50360

	Minimum wavelength	**Typical wavelength**	**Maximal wavelength**	**Unit**
Red light	618	625	630	nm
Green light	520	525	535	nm
Blue light	455	460	465	nm

Table 4.1: high-power full color LEDs (F50360) from Seoul Semiconductor

LEDs were installed on a designed PCB (Appendix *LED PCBs*). The light intensity of the diodes is controlled by a waveform generator, as shown in Figure 4.1. Such solution enables to vary different colors ratios of emitted light in order to measure a pixel sensitivity to change in colors ratios. The light from two separate RGB diodes is mixed and transmitted by a dual branch light guide (Dolan-Jenner E809). A light guide mount from a LED side and a custom PCB side was designed in *SolidWorks* and 3D printed (assembly drawing in (Appendix *Drawings in SolidWorks*). Measurement used to characterise stacked photodiode included a BPS101 Optical Feedback Controlled Tungsten Halogen Light Source and color

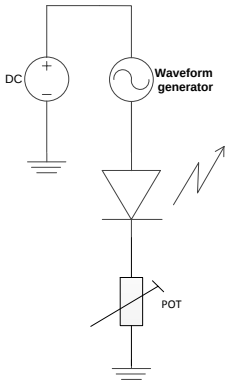


Fig. 4.1: Schematic of a LED driver.

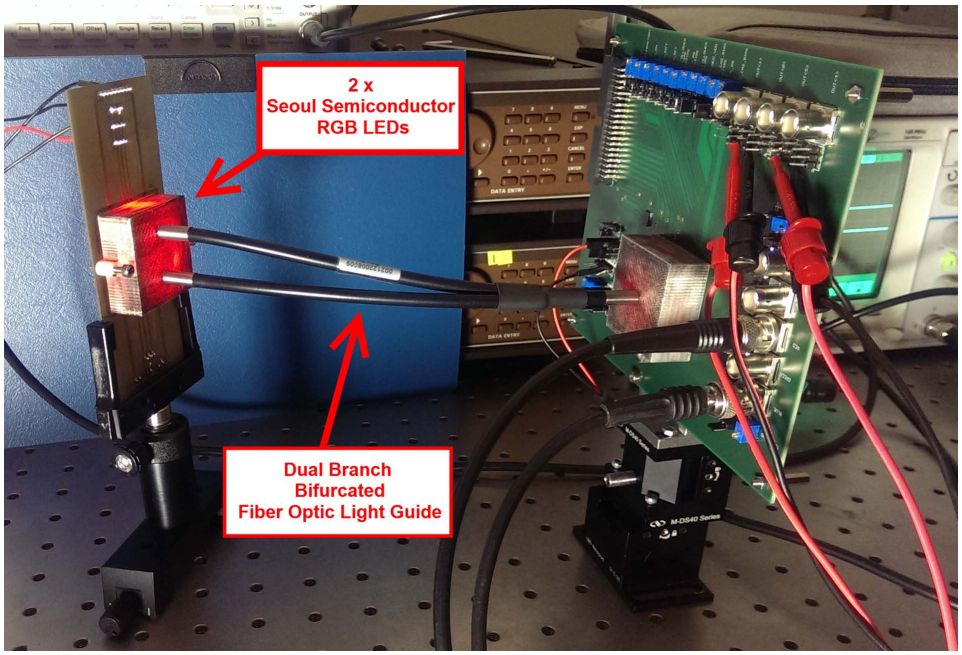


Fig. 4.2: Experimental setup to test a single pixel response and the photodiode. Light from RGB LEDs is transmitted by a fiber cable. The fiber cable is attached to PCBs by a 3D printed housing.

filters centred between $400nm$, $450nm$, $500nm$, $600nm$, $650nm$, $700nm$ and $880nm$. The filters centred at $450nm$ and $880nm$ have $80nm$ bandwidth and the other $50nm$. In order to measure a DC gain of the photoreceptors in 3C-Pixel neutral density filters with different light attenuations were used.

4.1.2 3C-TVS vision sensor measurement setup

In order to read out an image from a designed vision sensor a custom PCB was equipped with an IDC-40 bus realising an AER communication with the USBAERmini2 board [38], which is a part of an open source project jAER[39]. The USBAERmini2 board is equipped with a microcontroller handling the AER communication with the vision sensor and sending out the AER addresses to computer through a USB port. A software in JAVA language was written in order to display an image from a vision sensor. The vision sensor was stimulated by a beam of a collimated light (strong

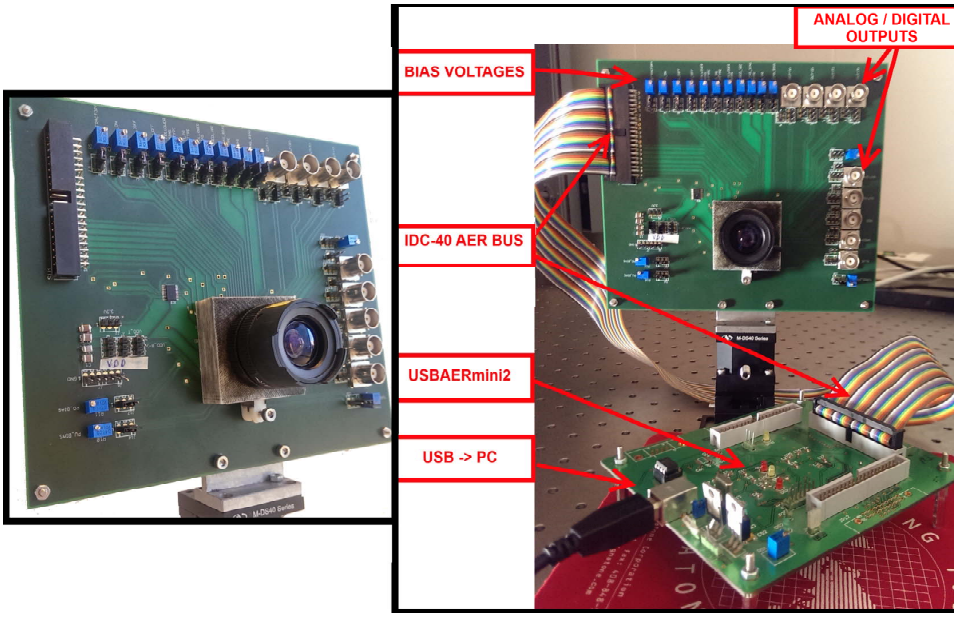


Fig. 4.3: Left photo: designed PCB with a $4mmF/1.2$ C-mount lens. Right photo: Designed PCB connected by a bus with the board USBAERmini2.

light dot) on the different background color. The background color was projected by another PCB equipped with six single color LEDs (2 x red - SZR05A0A, 2 x green - SZG05A0A, 2 x blue - SZB05A0A). Such solution provided a scattered light suitable to project a background light. The camera was tested with a $4mm$ F/1.2 C-mount lens installed on a designed in *SolidWorks* and 3D printed custom mount.

4.2 Spectral characterisation of the stacked photodiode

The stacked photodiode realises a front-end stage supplying three photocurrents: i_b, i_{b+g} and i_{r+g}. It is important to characterise a spectral sensitivity of the stacked photodiode in order to properly decode the ON/OFF events into colour changes. A surface of the implemented chip was illuminated by the BPS101 Tungsten Halogen Light Source with different colour filters centred between $400nm$ and $950nm$. For each colour filter, three photocurrents: i_b, i_{b+g} and i_{r+g} from a test photodiode were measured. The values were corrected by the Halogen Light Source irradiance factor as a function of wavelength, shown Figure 4.6a, and the colour filters transmission efficiency from Figure 4.6b. The reverse bias currents from the diodes used for the *ESD* protection of the padframe were not subtracted from the measured values of the photocurrents because we are only interested in ratios between the photocurrents and not the absolute values. The resulting spectral characterisation of the test photodiode is shown in Figure 4.5. Based on these results a relative light energies could be reconstructed. Figure 4.7a shows a relative response of each photodiode. The red photodiode is the photodiode located deepest in substrate, whereas the blue one is located on the surface of the substrate. Ideally, each photodiode should be sensitive only to the light of the corresponding colour result-

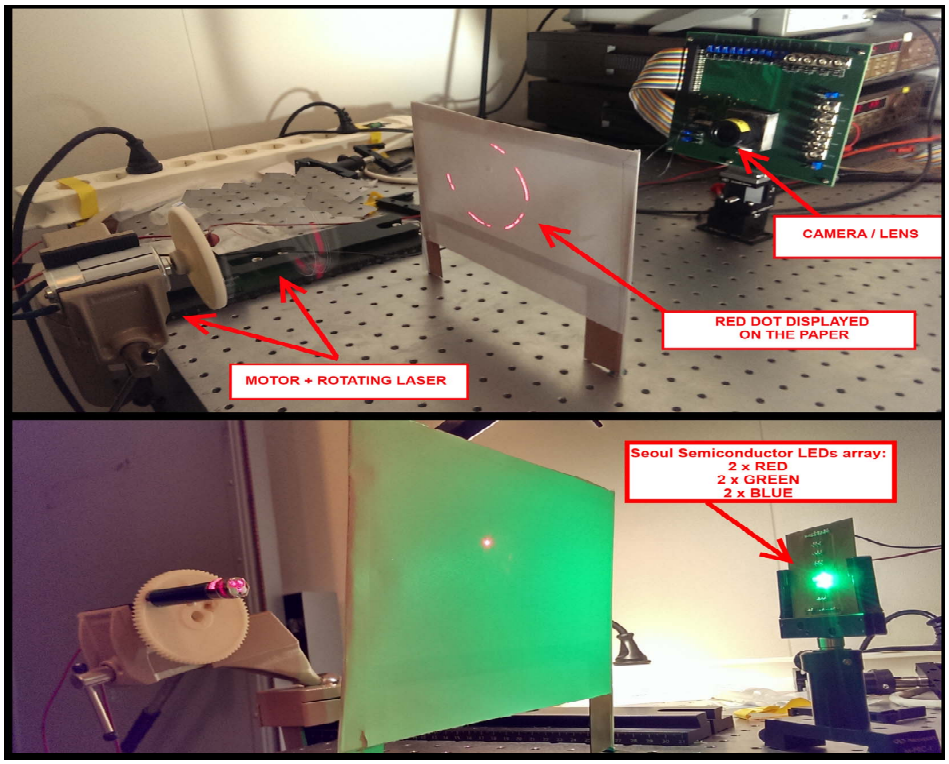


Fig. 4.4: Experimental setup to test a matrix response. A light source attached to the motor projects a fast rotating dot on the white sheet of paper.

ing in the photocurrents peaks around $620-750nm$ for the red photodiode, $495-570nm$ for the green photodiode and $450-495nm$ for the blue photodiode. Places where these peaks are expected are indicated by arrows. In fact we can observe that peaks are indeed located in the indicated places. The results from Figure 4.7a indicate that stacked photodiode has much higher sensitivity to long wavelengths rather than the short wavelengths from a visible spectrum. This behaviour may be explained by the structure of the photodiode: the size of the red photodiode is biggest and hence its depletion region is bigger than the other ones. As a result, the red photodiode has the biggest volume where the electric field close to the depletion region can catch electrons generated by the incident photons and therefore produce a photocurrent. Possibly the top photodiode PN junction and the middle PN junction are not located very deep, thus only very energetic photons can be caught by these photodiodes. This may explain worse sensitivity to the green and blue light. Another reason may be the fact that the biggest depletion area is obtained from the junction PSUB and DNWL because the dopants density in substrate is low and hence the depletion region is more extensive. Figure 4.7b shows relative responses of the stacked photodiode including mixtures of photocurrents in form of normalised $blue$, $red+green$ and $blue+green$ photocurrents. In order to be able to distinguish between colours, it is important to analyse photocurrents fractions: $(I_b)/(I_b+I_g)$, $(I_r+I_g)/(I_g+I_b)$ as a function of wavelength as shown in Figure 4.8. This plot is very important because it shows how the stacked photo diodes can be used to discriminate between colors: the $(I_r+I_g)/(I_g+I_b)$ ratio increases when a the dominating light wavelength increases (the color

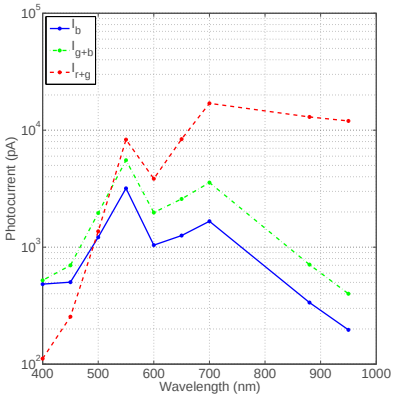
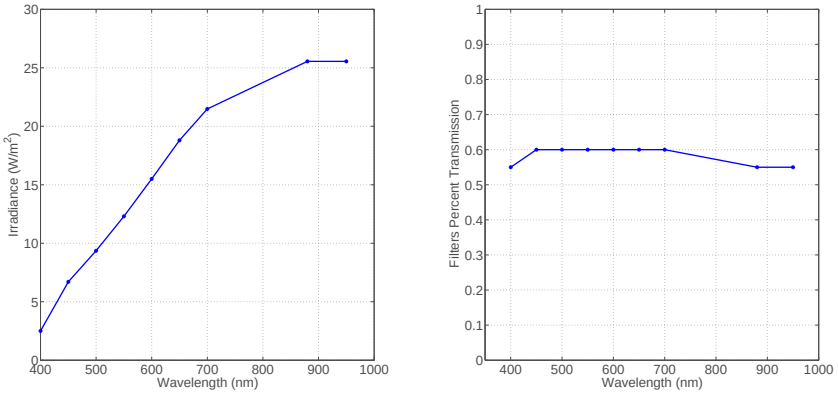


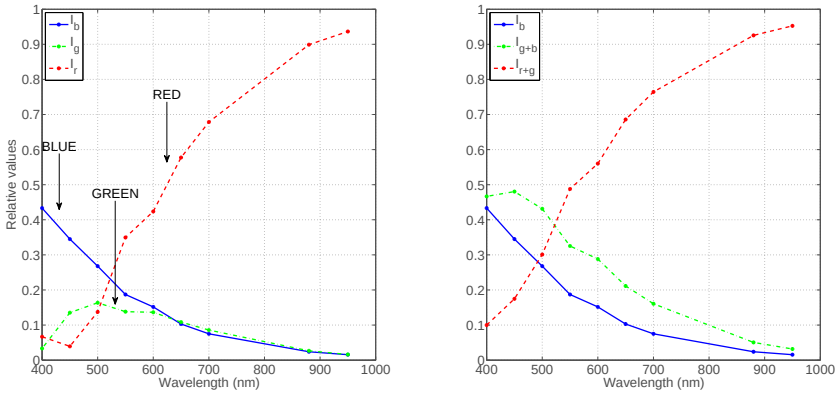
Fig. 4.5: Corrected diode photocurrents



(a) White Source Irradiance (b) Filters transmission efficiency

Fig. 4.6: Halogen Light Source irradiance and the colour filters transmission efficiency correction factors.

spectrum becomes *redder*) and the ratio $(I_b)/(I_b + I_g)$ decreases when the dominating light wavelength decreases (the color spectrum becomes *bluer*). The ratio $(I_b)/(I_b + I_g)$ exhibits weaker dependence on color spectra, it has steepest slope up to $650nm$ and after that is flat, while the $(I_r + I_g)/(I_g + I_b)$ ratio has almost constant slope over the entire tested spectrum. Thus if for instance a narrow band light source moves from $600nm$ to $900nm$ wavelength, it would lead to positive $(I_r + I_g)/(I_g + I_b)$ events and a few negative $(I_b)/(I_b + I_g)$ events. These observations are consistent with Table 3.7 from the previous chapter. This Table is repeated here for the convenience below. According to this table, we simplify and assume the photocurrents I_{r+g}, I_{b+g} and I_b to represent red, green, and blue spectra respectively. As a result, the photocurrents ratios become R/G and B/G and their changes can be now translated into *bluer*, *greener* or *redder* color change. Finally, the ON1 events is assumed to be translated into a blue colour increase because the fraction $\frac{I_b}{I_b+I_g}$ increase is *strongest* for a blue light. The OFF2 and OFF1 events are assumed to be translated into an increase of a green colour. The ON2 event will be translated into a red light change increase.
In order to compare experimental results with a simulation, the coefficients describing behaviour of the stacked photodiode are extracted from Fig-



(a) Relative responses of single photodiodes in the stacked photodiode.

(b) Relative responses of the stacked photodiode.

Fig. 4.7: Spectral response of a stacked photodiode in TSMC 90nm process.

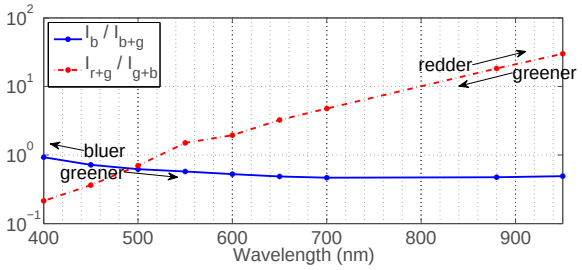


Fig. 4.8: Photo currents ratios amplified by the differencing amplifiers (Equations 3.30,3.31). The values were corrected by the Halogen Light Source irradiance factor and color filters transmission efficiency.

ure 4.7a. These coefficients indicate how much photocurrent is obtained from each photodiode when illuminated by different light colors. Further, these coefficients will be used to emulate the actual stacked photodiode while simulating the 3C-Pixel and compare the simulation results with the experimental results. The coefficients are gathered in Table 4.3.

	Red photodiode	**Green photodiode**	**Blue photodiode**
Red light 625nm	0.68	0.16	0.16
Green light 525nm	0.47	0.22	0.31
Blue light 460nm	0.175	0.275	0.55

Table 4.3: Extracted coefficients of the photocurrents for the stacked photodiode, when illuminated by different light colors: red light - 625nm, green light - 525nm, blue light - 460nm

Event	Photocurrents ratio	Assumed spectra ratio	Color change
ON1	$i_b/i_{b+g}\nearrow$	$\propto \frac{B}{G}\nearrow$	bluer
OFF1	$i_b/i_{b+g}\searrow$	$\propto \frac{B}{G}\searrow$	greener
ON2	$i_{r+g}/i_{b+g}\nearrow$	$\propto \frac{R}{G}\nearrow$	redder
OFF2	$i_{r+g}/i_{b+g}\searrow$	$\propto \frac{R}{G}\searrow$	greener

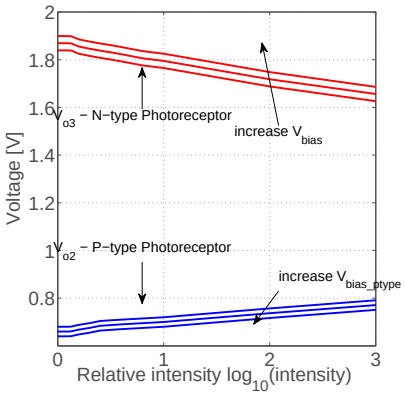
Table 4.2: Translation of ONx, OFFx events into color changes. The density of ONx/OFFx events is proportional to the rate of the color changes.

4.3 Test 3C-Pixel

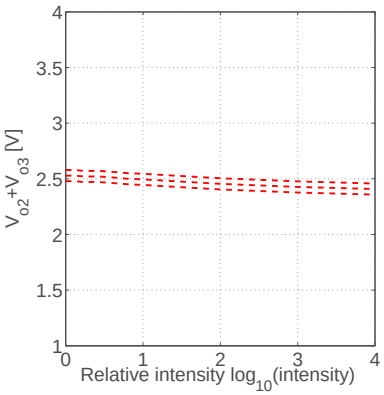
4.3.1 Logarithmic photoreceptors.

DC response

In order to test if the complementary photoreceptors have the same gain, their DC responses was measured. The surface of the camera was illuminated by the coherent flat spectrum light source. Different



(a) Response of an N-type Photoreceptor - V_{o3}, and P-type Photoreceptor - V_{o2}



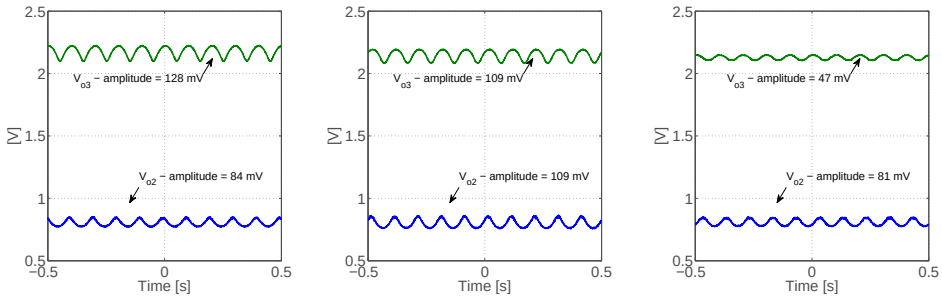
(b) Sum of V_{o2} and V_{o3} as function of for different bias voltages as function oflight intensity.
light intensity.

Fig. 4.9: DC responses of the complementary Photoreceptors - experimental results.

attenuating filters were placed in front of the chip in order to sweep the light intensity. The resulting responses of the logarithmic photoreceptors from the test pixel are shown in Figure 4.9a. Ideally, the sum of these two responses should give a flat line. Figure 4.9b shows that both photoreceptors have very similar DC gain. It is impossible to compare the DC gain from simulations from Figure 3.7 with these experimental results because we do not know the absolute values of the photoccurents from the test diodes (we know only the relative changes). Since the photoreceptors are logarithmic amplifiers, their responses on the logarithmic x-axis scale give a straight line what is also shown in Figure 4.9a.

Transient response

The surface of the chip was illuminated by a sinusoidal light of different colors and the responses of logarithmic photoreceptors were registered. Figure 4.10 shows results for a red, green and blue light stimuli. The



(a) Red (b) Green (c) Blue

Fig. 4.10: The output of the photoreceptors: V_{o2} and V_{o3} when illuminated by different light colors: red light - 625nm, green light - 525nm, blue light - 460nm

results are consistent with the relative responses of the stacked photodiodes from Figure 4.7b. For the red stimuli the strongest response is achieved from the photoreceptor V_{o3} (Figure 4.10a). For a green light (525nm) both responses are identical (Figure 4.10b) and for a blue light (460nm) V_{o2} has the response with the biggest amplitude (Figure 4.10c).

4.3.2 Pixel bandwidth

The test pixel was stimulated by a sinusoidal red LED light of approximately 1klux brightness on the white light background. The number of events

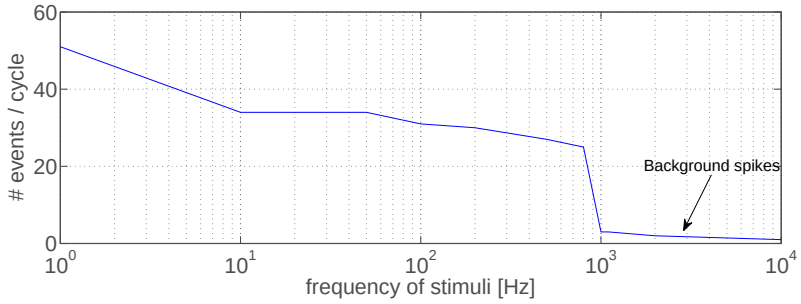


Fig. 4.11: Number of recorded events per cycle as a function of the stimuli frequency. The stimuli was a sinusoidal red LED light of 1klux brightness.

was counted for different frequency of this stimuli. The measurement was conducted for 20 cycles of stimuli and averaged. The resulting frequency response is shown in Figure 4.11. Around $800Hz$ the pixel does not respond any more to the stimuli and we registered only background events. Probably the capacitance from the differencing amplifier limited the pixel bandwidth.

4.3.3 Color change representation

The designed pixel monitors three different color spectra and it generates four different events whenever the ratio of the light intensity filtered by those three spectra changes. Based on the spectral characterisation of the stacked photodiodes we decided to assign events to the colour changes according to Table 4.2. Now we will illuminate the chip surface by two sinusoidal out-of-face different colour LED light sources. In this way we will test the pixel response to three different colour opponencies: R/G, B/G and R/B and we will compare those resposes to the pre-layout and post-layout simulations. Generally, we observed higher sensitivity of the pixel

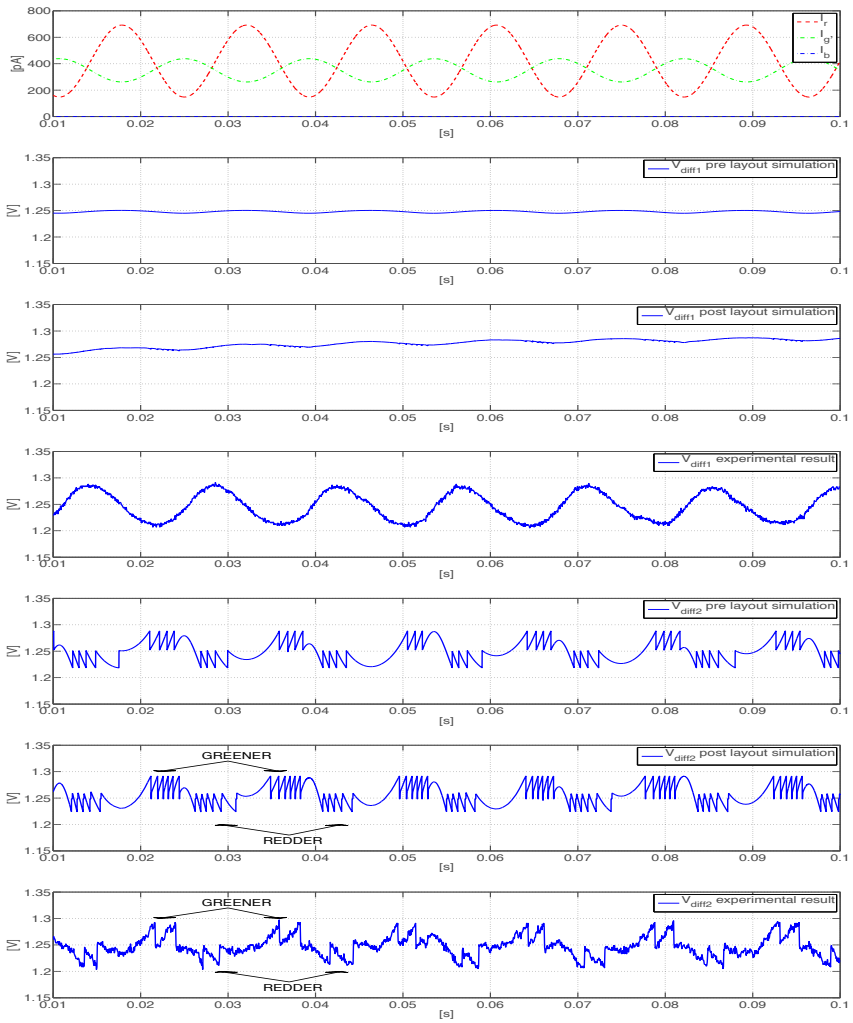


Fig. 4.12: A single pixel response to the continuous *red-green* contrast change. From top: modelled input photocurrents, V_{diff1} pre layout transient simulation, V_{diff1} post layout transient simulation, V_{diff1} experimental result,V_{diff2} pre layout transient simulation, V_{diff2} post layout transient simulation, V_{diff2} experimental result.

in the post-layout simulations, meaning that the differencing amplifiers have higher gain. Higher gain can be explained by higher capacitance from C_3, C_4, C_5, C_6 because of the parasitic capacitors. Under the C_3, C_4, C_5, C_6 Capacitor-Bottom-Metal (CBM) layer there is a long wide metal M7 routed which is connected to the Capacitor-Top-Metal (CTM) layer. Such

metal organisation created so called *sandwich* structure CTM-CBM-CTM which increased the overals capacitance of C_3, C_4, C_5, C_6 and resulted in higher differencing amplifier gain. On the other hand, number of spikes in the experimental results differs from the results in simulations because we do not know the absolute values of the photoccurents from the test photodiodes (we know only the relative changes).

Red-Green Contrast Change

First, we characterised the pixel response to the continuous *red-green* intensity change. The top plot in Figure 4.12 shows the modelled input photocurrents and the other plots show pixel (pre-layout, post-layout and experimental) response to such stimuli. The thresholds used during simulations were: V_{ON} = 1.29 and V_{OFF} = 1.22. Coefficients from Table 4.3 were used in order to emulate the actual stacked photodiode response while simulating the 3C-Pixel and compare the simulation results with the experimental results. The experimental results were obtained by illuminating the chip surface with a sinusoidal out-of-phase red and green light sources of the similar frequency used in simulations. The thresholds for the tested chip were: V_{ON} = 1.3 and V_{OFF} = 1.2. The transient plots for experimental results of V_{diffx} suggest that the actual thresholds were lower. However, it has to be taken into consideration that the output voltages are buffered by non ideal analog buffers and the scope resolution was not high enough to track *peaks* of V_{diffx} right before the reset signal. For the *red-green* intensity change stimuli we observed only ON2 and OFF2 events. For the pre layout and the post layout simulations we observed 4-6 events per a sinusoid cycle, while for the experimental results 2 per a sinusoid cycle. The difference between experimental results and simulations can be explained by the fact that actually we do not know the real values of photocurrent generated by the stacked photodiodes implemented in silicon. Since we were changing the red and green spectra inversely, the change of the R/G ratio is most substantial and hence the output V_{diff2} exceeds the V_{ON2} and V_{OFF2} thresholds. V_{diff1} moves between the thresholds but it does not exceed them, and hence, no ON1 nor OFF1 events are generated. Since there has not been any blue light change in the stimulus (the blue LED was off), the nominator and denominator in the B/G ratio change by the same amount, and ideally this ratio should remain constant. In fact, the output V_{diff1} fluctuates only because the photo currents ratio B/G exhibits weak dependence on longer spectrum wavelengths, as was shown in Figure 4.8.

Blue-Red Contrast Change

Next, we characterised in the same way the pixel response to the continuous *blue-red* intensity change. Figure 4.13 shows the modelled input photocurrents and the pixel (pre-layout and post-layout) response to such stimuli. In this case we observed four types of events: events ON1 occur when the blue light positive transient is biggest (B/G increases) and conversely, ON2 events occur at the maximum red positive transient (R/G increases). Besides, we observed OFF1 and OFF2 events that might be confusing at first glance. However, OFF1 events indicate a shift to longer wavelengths from blue and OFF2 events indicate a shift to shorter wavelengths from red and

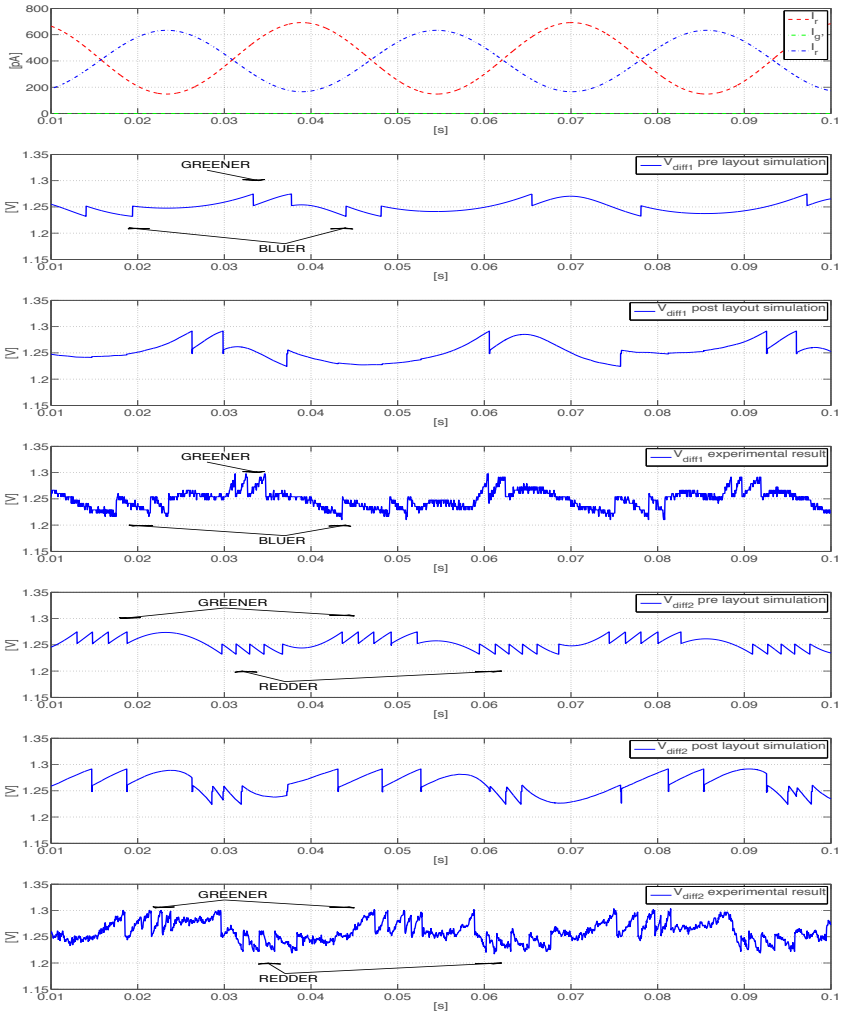


Fig. 4.13: A single pixel response to the continuous *red-blue* contrast change. From top: modelled input photocurrents, V_{diff1} pre layout transient simulation, V_{diff1} post layout transient simulation, V_{diff1} experimental result, V_{diff2} pre layout transient simulation, V_{diff2} post layout transient simulation, V_{diff2} experimental result.

thus, the responses are as expected and correct. We observed more ON2 and OFF2 events because, as shown in Figure 4.8, the R/G ratio has steeper slope. As a result, if the light spectra is changing towards longer wavelengths (*redder*), we should observe more ON2 and OFF2 events.
We observed that V_{diff1} slowly drifts towards V_{OFF} threshold what provokes unwanted events. The explanation of such behaviour may be an issue with wrong layout and a parasitic capacitance at the input of the amplifier. This problem applied only to the single test pixel, not pixels in the array (tested later). The test pixel has probes connected to its sensitive nodes, and this could lead to the increase of parasitic capacitances and make the source follower to charge the input of the differencing amplifier and increase its output, resulting in unwanted events.

Blue-Green Contrast Change

Finally, we characterised the pixel response to the continuous *blue-green* intensity change. Figure 4.14 shows the results. Directly responsible for

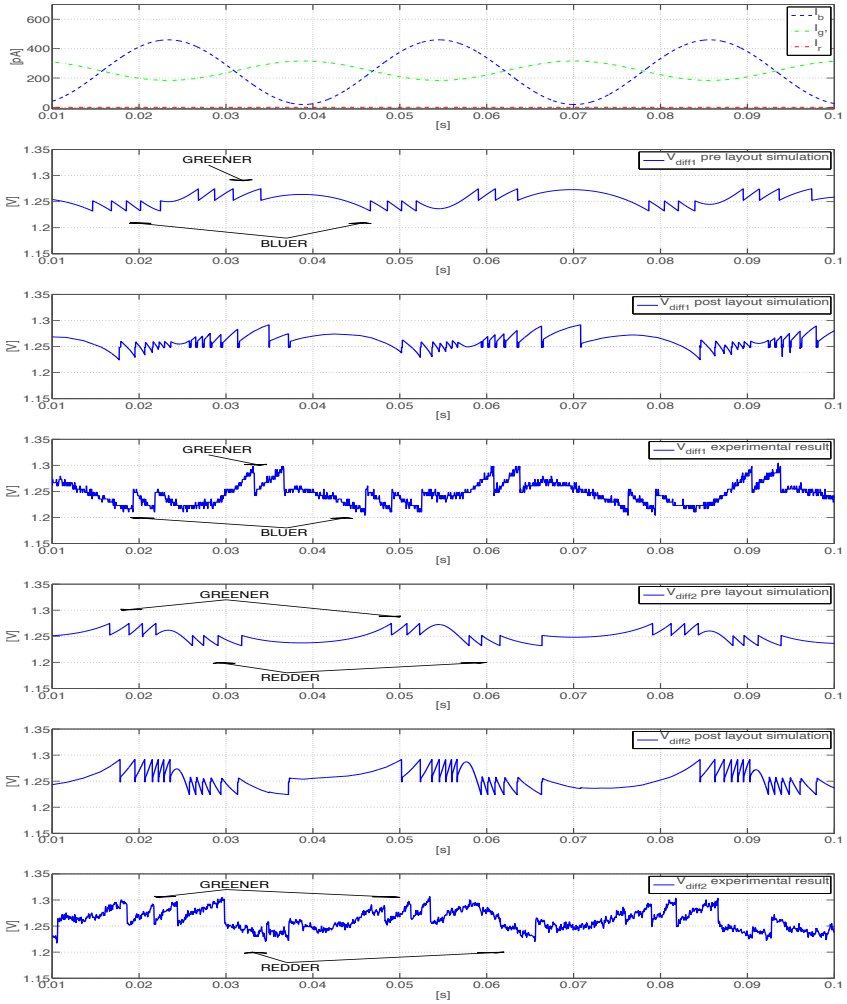


Fig. 4.14: A single pixel response to the continuous *blue-green* contrast change. From top: modelled input photocurrents, V_{diff1} pre layout transient simulation, V_{diff1} post layout transient simulation, V_{diff1} experimental result, V_{diff2} pre layout transient simulation, V_{diff2} post layout transient simulation, V_{diff2} experimental result.

sensing the *blue-green* contrast change are events ON1 and OFF1. Indeed, out of all three different colour opponencies tested, the B/G opponency results in the pixel strongest response in terms of ON1 and OFF1 events. We observe OFF1 events for the positive slope of I_g, and conversely, ON1 events for the positive slope of I_b. This same applies to the experimental results. Regarding the second R/G opponency we observed ON2 and OFF2 events that again might be confusing. However, we should recall that we simplified the ratio $I_r + I_g / I_g + I_b$ into I_r / I_g as shown in Table 4.2. In the denominator of the ratio $I_r + I_g / I_g + I_b$ there is actually a photocurrent from a blue spectra. Hence for the positive slope of I_b stimuli we observe OFF2 events and for negative slope of I_b stimuli we observe ON2 events.

4.3.4 Intensity change

The differencing amplifiers amplify photo currents ratios according to Equations 3.30,3.31). If there is only one narrow band light source changing its intensity, the nominator and the denominator should change proportionally and the the color change ratios should remain constant. Hence, we should not observe any events. Figure 4.15 shows simulation

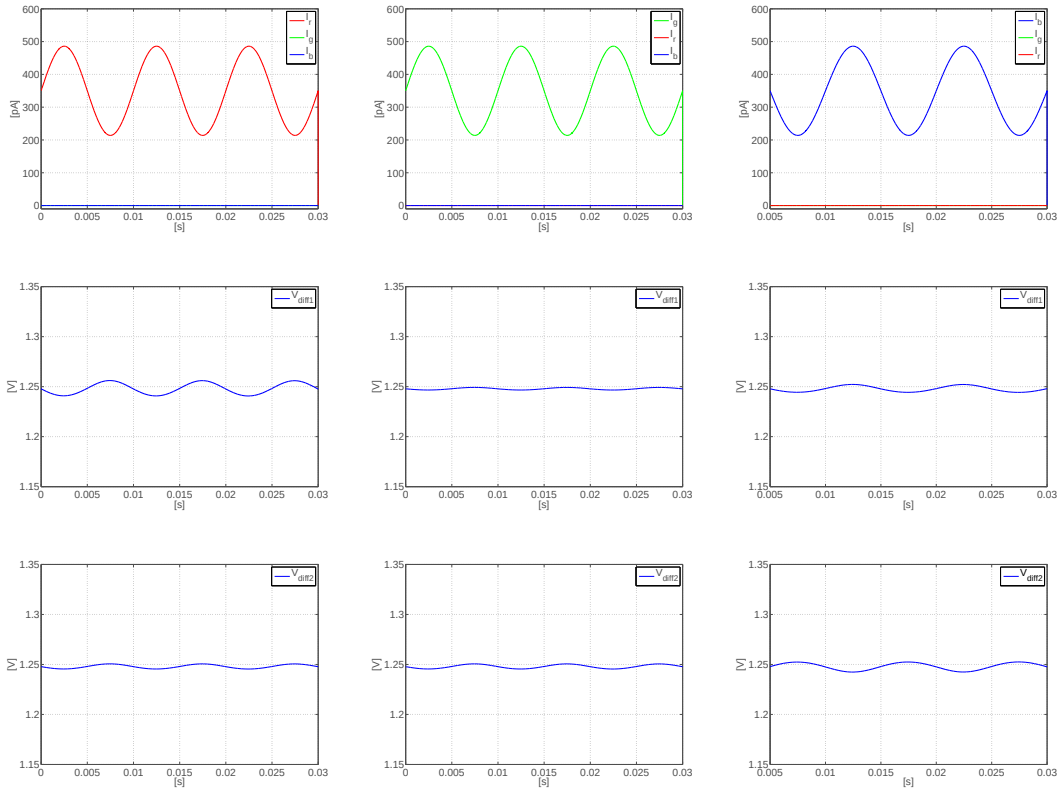


Fig. 4.15: Intensity change simulation. The pixel is stimulated by only one photocurrent I_r, I_g or I_b at the time.

results when the pixel is stimulated by only one sinusoidal photocurrent I_r, I_g or I_b at the time. For all three cases V_{diffx} fluctuate between threshold values set between $1.2V - 1.3V$. Figure 4.16 shows the same test, but performed on the chip. The test pixel was illuminated by only one color LED light (R, G or B) at the time with no background light. The first three top plots show the current consumption of the LEDs. As it was in simulations, V_{diffx} fluctuate and does not exceed the threshold values set between $1.2V - 1.3V$. The outputs are noisy because, in this case even few random photons of different wavelength which incident on the chip can make both ratios vary a lot.

4.3.5 Close look on two events

Figure 4.17 shows a close look on V_{diff2} and RESET signals. During the RESET signal is low V_{diff2} is roughly $1.22V$. When the RESET signal goes high, V_{diff2} rises almost $2mV$ because of the charge injection and the transistor feedthrough. During switching the reset transistor (M22 and M26 in Figure 3.20), an unwanted negative charge from this transistor

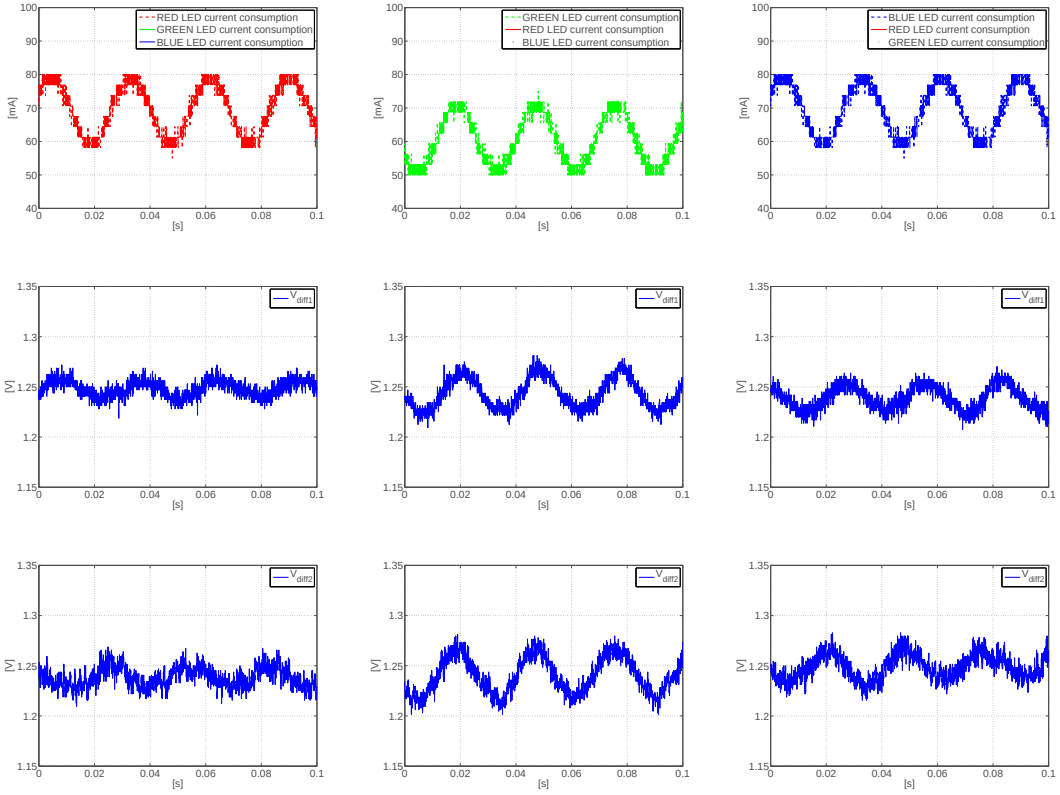


Fig. 4.16: Intensity change test - experimental results. The test pixel was illuminated by only one color LED light (R, G or B) at the time in absolutely dark light conditions.

channel is injected into the inverting input of the amplifier. Also, the gate-source capacitance of this reset transistor *couples* the RESET and V_{diff2} signals. For high speed switching RESET signal from GND to VDD this gate-source capacitance acts temporarily as a short. As a result, output V_{diff2} moves up after balancing. This results in statistical offset and higher sensitivity of the pixel to OFF events. This is not a substantial problem, because this offset can be compensated by proper adjusting thresholds.

4.3.6 Refractory time functionality

Figure 4.18 shows that the number of spikes increases when the V_{refr} decreases. The bias voltage V_{refr} controls the refractory time of the pixel, by changing the conductivity of the pull-up transistor equivalent to the refractory time of the neuron. The chip was illuminated by a constant *green-red* color change stimuli and number of spikes for different V_{refr} was counted. For instance, for $V_{refr} = 2.12V$ we observed only 2 events per cycle and the refractory time was approximately $14ms$. For $V_{refr} = 1.62V$ we observed roughly 12 events per cycle with the refractory time below $100\mu s$.

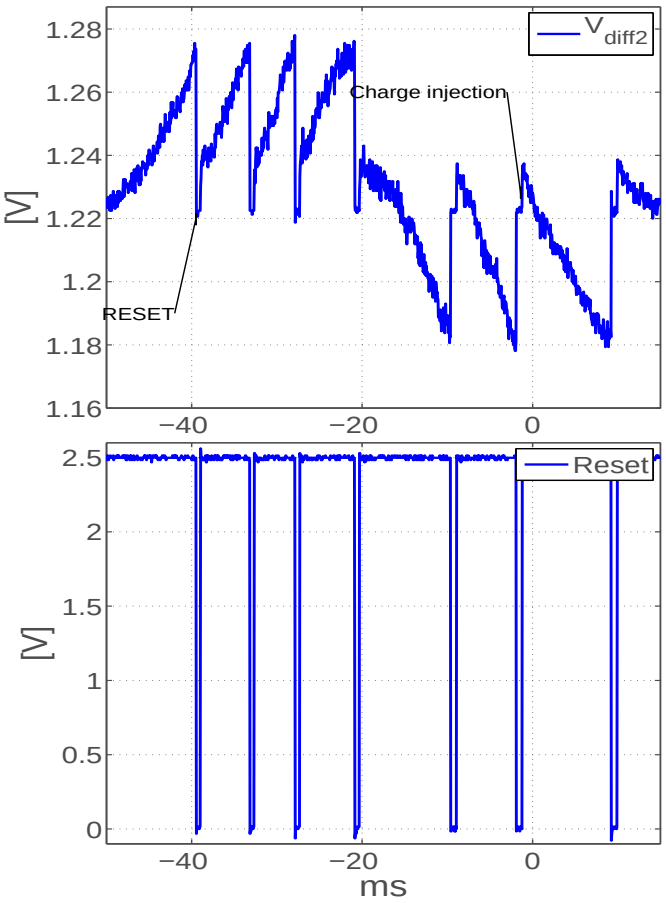
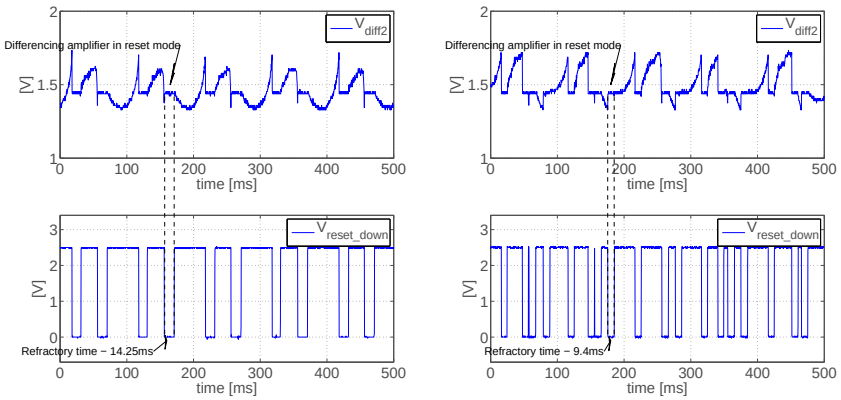


Fig. 4.17: Close look on V_{diff2} and RESET signals. During the RESET signal is low V_{diff2} is roughly $1.22V$. When the RESET signal goes high, V_{diff2} rises almost $2mV$ because of the charge injection and the transistor feedthrough.

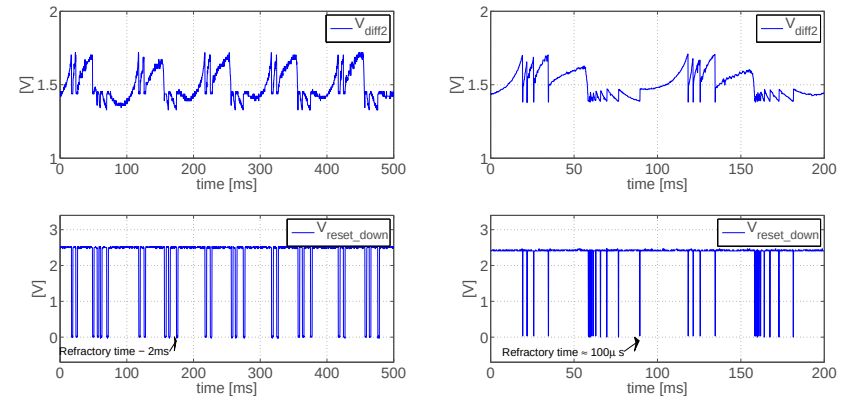
4.4 Vision sensor

4.4.1 Fixed-Pattern Noise

As for the traditional cameras, we want to test a Fixed-Pattern Noise (*FPN*) figure of the designed vision sensor. For the frame-based cameras *FPN* describes if for the constant illumination level, all pixels have the same output levels or some pixels respond stronger or weaker than the rest. For the temporal contrast vision sensor the same test was performed with the flat spectrum light source set in front of the vision senor. All possible sources of unwanted noise, as flickering light bulbs, computer displays, were turned off. Such test should provide uniform scene conditions for all pixels. Number of ON1 and OFF1 events from each pixel was counted. The measurement lasted 3 minutes for the background light intensity of $100lux$. Figure 4.19 shows number of ON1 and OFF1 spikes recorded by each pixel. The first and the last row of the pixels was not included in the test because those pixels were located close to the AER peripheries and the test buffers, what increased their mismatch. Mean number of ON1 events is 22, standard deviation: 28 and mean number of OFF1 events is 18.8 and standard deviation 24. The Fixed-Pattern Noise was counted as a fraction



(a) V_refr=2.12V. 2 events per cycle. (b) V_refr=2.11V. 4 events per cycle.



(c) V_refr=2.01V. ≈ 6 events per cycle. (d) V_refr=1.62V. ≈ 12 events per cycle.

Fig. 4.18: V_diff2 and V_reset_down transient plot for different V_refr - test of a refractory time change functionality.

between a standard deviation of all pixels and the average number of spikes:

$$FPN = \frac{\sigma_{x,y}}{\hat{n}} \tag{4.1}$$

Fixed-Pattern Noise for ON1 and OFF1 is::

$$FPN_{OFF1} = 78\%. \tag{4.2}$$

$$FPN_{ON1} = 78\% \tag{4.3}$$

The Fixed-Pattern Noise is very high for the static $100lux$ white light stimuli because few pixels spike with much higher frequency then the average. We observed that FPN improves proportionally to the background light intensity and the average spiking rate decreases if the background light increases.

4.4.2 Uniformity of the response

The next test is similar to the previous one, but this time we want to test a uniformity of the pixels response to a slow (below $1Hz$) *blue-green* background light change. We used again the measurement setup shown in

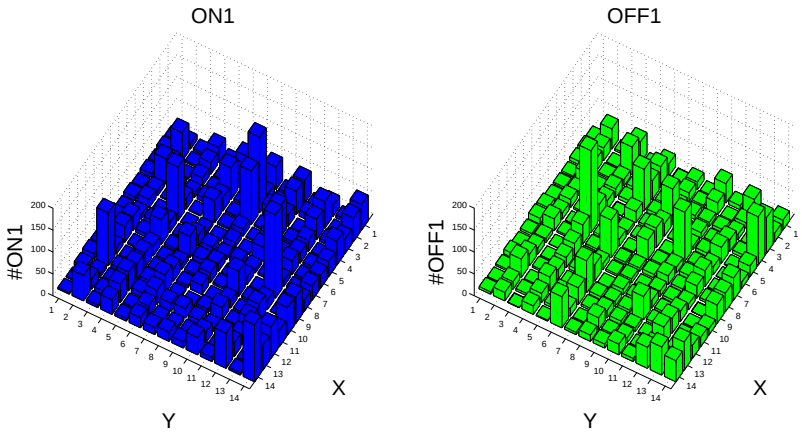


Fig. 4.19: Response of the camera to the static flat spectrum background light measured during 3 minutes.

Figure 4.3 and we did not project any pattern on the sheet of paper, but we slowly changed the background light from blue to green and vice versa, and counted number of ON1 and OFF1 spikes from each pixel. Such test should provide slow and uniform contrast change conditions for all pixels. The measurement was repeated 20 times and the results were normalised to one background light change. Figure 4.20a shows number of spikes recorded per one slow *blue-green* background light change. The visual inspection suggests that the pixels have uniform response.
Figure 4.20b shows a distribution of events recorded per one *blue-green* background change. The shape of the events distribution is similar to the Gaussian distribution function. The Fixed-Pattern Noise for ON1 and OFF1 for a slow background light change is:

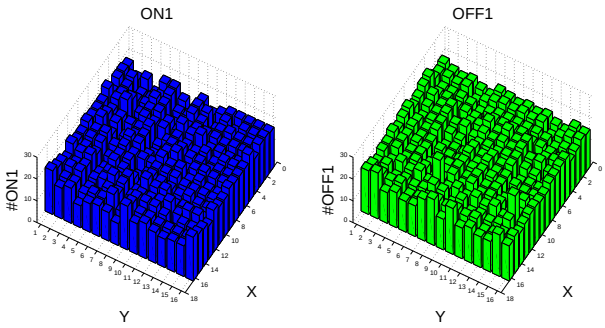
$$FPN_{OFF1} = 39\%. \tag{4.4}$$

$$FPN_{ON1} = 30\% \tag{4.5}$$

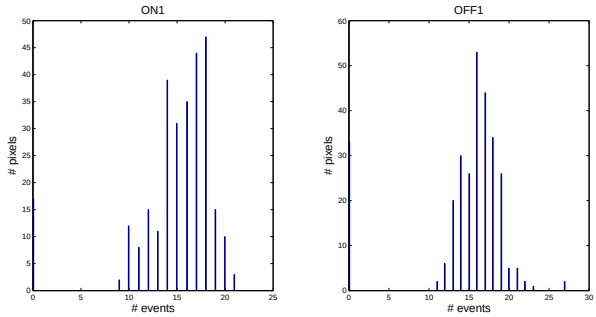
For this test we obtained better FPN than for a static white light.

Two colours change stimuli

Figure 4.21 shows the response of the vision sensor to a rotating red dot on a green background and Figure 4.22 shows the response of the vision sensor to a rotating blue dot on a red background. According to Table 4.2 we assigned green color to OFF1 and OFF2 events, red color to ON2 events, and blue color to ON1 events. Comparing Figures 4.21 and 4.22 we can observe that for a rotating red dot roughly 5 times more events representing change of R/G ratio are registered. For a rotating blue dot the vision sensor generates much less events representing change of R/G ratio. In Figure.4.21 the red events preceded the center of the red dot stimulus while the green and blue events are trailing. On the other hand, in Figure 4.22 the blue and green events preceed the center of the red dot, while red events are trailing.



(a) Normalised respone of the matrix to the slowly changing *blue-green* background light. Entire number of ON1 events: 868, OFF1 events: 842.



(b) Histogram showing a distribution of events ON1 and OFF1 recorded per one *blue-green* slow background change. Mean number of ON1 events: 14.7 and standard deviation: 4.4. Mean number of OFF1 events 14.5 and standard deviation: 5.7.

Fig. 4.20: Matrix uniformity test.

White light stimuli

Figure 4.23 shows the response of the vision sensor to a rotating white dot on a white background. Ideally, if we change only white light intensity, both ratios amplified inside the pixels should remain constant and we should not register any events. However, for this experiment we observed events. The same behavior was observed in [18], for a white light stimuli, a leading edge appeared red and a blue-green tail was trailing. Probably, the reason of such behavior is the difference between the speed of the stacked photodiodes. Each junction differs in thickness of the depletion regions, and hence in parasitic capacitances of these junctions. As a result, both ratios amplified inside the pixels do not remain constant for a flat spectrum light stimuli. Another reason of observing spikes may be wrong setup: for a background light we used a work-lamp fluorescent lamps when the rotating dot was created by a coherent flat spectrum light source. Probably spectra of these two light sources was not identical.

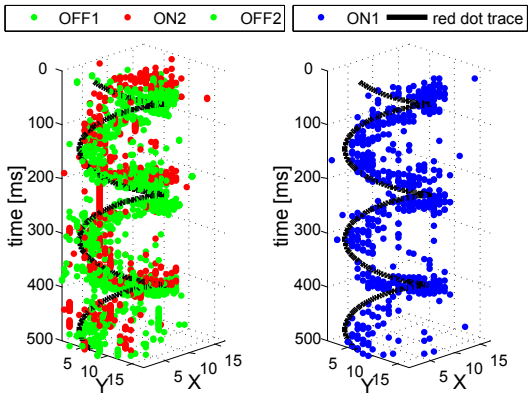


Fig. 4.21: Space-time plot of the rotating red dot on the green background. Number of events: ON1=924, OFF1=1107, ON2=5060, OFF2=5210.

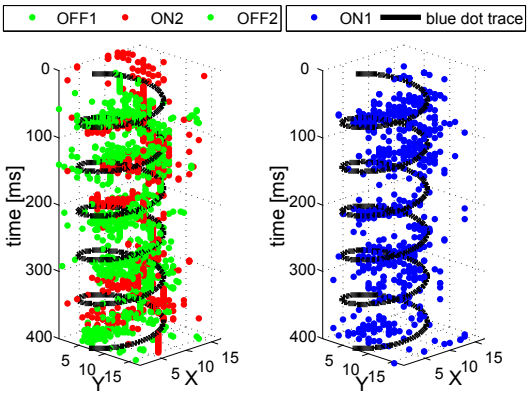


Fig. 4.22: Space-time plot of the rotating blue dot on the red background. Number of events: ON1=1208, OFF1=1353, ON2=2845, OFF2=2795.

4.4.3 Monochromatic image

We tried to use the camera to display monochromatic image, by showing all events as separate pixels. Figure 4.24 shows outputs of the vision sensor for different scenes. Figures 4.24a and 4.24b show the camera response for a red LED moving in front of the camera for different background light conditions. After a closer inspection we can distinguish ONx and OFFx event. ONx events are localised at rows with odd addresses and OFFx events are localised at rows with even addresses. We can easily distinguish *head* and *tail* of the moving stimuli. The *head* and the *tail* are the places with a high temporal color contrast between the diode and the background. The *head* and the *tail* are smaller for the image in Figure 4.24b because this image was taken with much stronger background light illumination ($15klux$). As a result the glow from the LED is smaller and the temporal color contrast is registered closer to the center of the LED. The light illumination in these images varied from $0.1lux$ up to $15klux$, suggesting that the pixel has a dynamic range equal at least 5 decades. Figure 4.24c shows an image of a fast rotating led dot with a white background light. These figures show that the camera can be also used to sense a temporal contrast in one color.

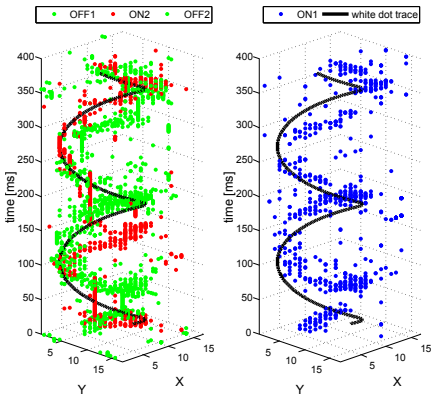
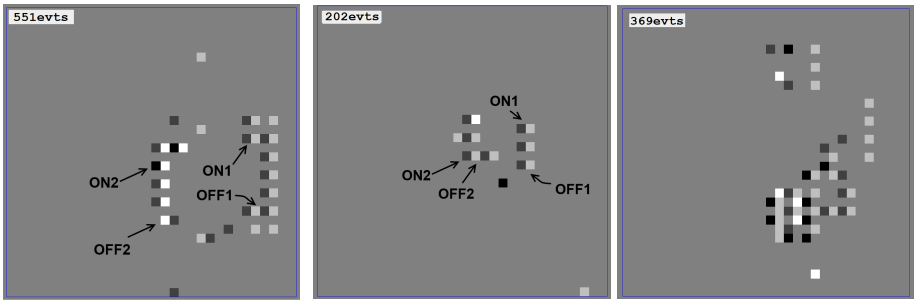


Fig. 4.23: Space-time plot of the rotating white dot on the white background. Number of events: ON1=2537, OFF1=4376, ON2=4030, OFF2=4058.



(a) Moving red LED from the right to left side. White background light illumination $0.1lux$

(b) Moving red LED from the right to left side. White background light illumination $15klux$

(c) Clockwise direction rotating red dot on the white background. Visible *head* and lagging *tail*.

Fig. 4.24: Monochromatic images recorded from the camera.

4.4.4 Color image

Figure 4.25 shows a color image obtained by the vision sensor. We customised a JAVA program from the open source project jAER[39] in order to display color changes according to Table 3.7. Red square was moved from right to left in front of the camera and the background light was set green. We can easily distinguish a red *head* of the rectangle which indicated the movement direction. The *head* indicates the place where the color contrast changes towards red spectrum. To the right from the *head* we can observe a green *tail*, which indicates the end of the rectangle. The*tail* is green because when the red stimuli moves away, the red spectrum decreases, and hence it is translated into increase of the green spectrum. We can observe *blue events* because the photocurrents ratio $(I_b)/(I_b + I_g)$ from the stacked photodiodes exhibits a weak dependence on red light.

4.4.5 Greediness of the arbiters

The arbitration implemented in this work is a full arbitration with a 'Greedy' Arbiter [32] shown in Figure 2.16 and discussed in chapter *Background*. The disadvantage of such arbitration is the greediness. The arbiters can continuously favour the same requests if the arbiters are

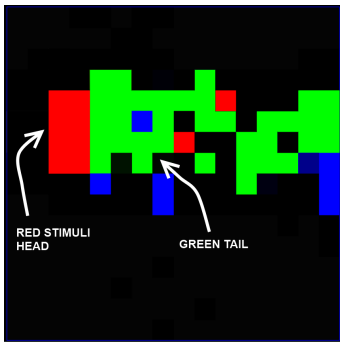


Fig. 4.25: Red stimuli on green background.

overrun. We tried to provoke such behaviour by the same measurement setup used to test a uniformity of response. This time the rate of the background color changes was much higher (over $200Hz$). In this case all pixel should generate multiple events at the same time, such that the arbiters should become overrun. The resulting events distribution of ONx and OFFx events is shown in Figure 4.26.

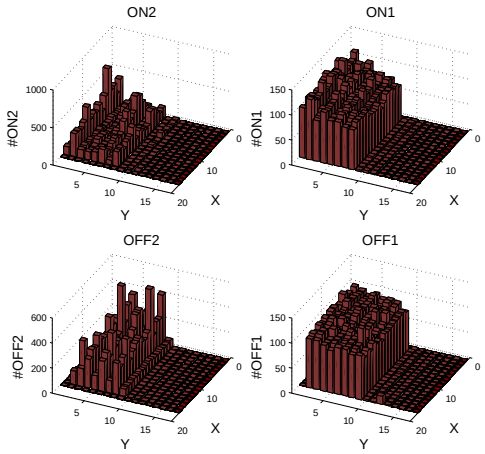


Fig. 4.26: Events distribution of the vision sensor response to fast background color changes.

Only half of the pixels received an acknowledge. We can easily guess that the most right Y axis arbiter in the arbiters tree, shown in Figure 2.18, favoures the top 8 pixels when it is flooded by a series of multiple consecutive requests.

4.4.6 Power consumption

The power consumption of the chip (excluding analog and digital buffers of the test pixel) varied between $19 \div 29mA$. The power consumption depends on comparators thresholds V_{ON} and V_{OFF} from Figure 3.17. These thresholds control number of spikes generated by the sensor, however the power consumption does not depend on the number of spikes generated. The power consumption increases when V_{ON} or V_{OFF} increases because the comparators are realized by common-source amplifiers and the amount of current they consume is proportional to these thresholds. The power consumption of a single pixel for pre-layout and post-layout simulations

was approximately $90\mu A$. Multiplying this value times the number of pixels gives entire power consumption of $23mA$, which is similar to the experimental measurement.

4.4.7 Substrate Coupling

We observed that the voltage representing the ratio between the photocurrents from the photodiodes located deepest in the substrate is distorted by noise. The main reason of this noise is the fact that the photodiode located deepest in the substrate is formed by a PN junction between *deep nwell* and the substrate. The substrate is not isolated but shared between other photodiodes. As a result, any current injected from an adjacent circuitry (especially digital circuitry) can be easily *coupled* by the substrate to a very sensitive photodiodes and amplified by the logarithmic photoreceptor. This coupling lead to a *cross-talk* between two channels inside the pixel and between the adjacent pixels. We observed a correlation between spikes, meaning that there was a high possibility to observe spikes $xxx1$ and $xxx2$ consecutively.

4.4.8 Instability issues

The photoreceptor passed tests in terms of stability when it was simulated. However, for very low photocurrents the photoreceptors in the actual 3C-Pixel start oscillating within approximately 100mV, what generates random ON/OFF events on the output of the pixel. Interestingly, according to the Figure 4.8 the phase margin for small photocurrents is very big, and the the photoreceptor is not expected to oscillate Even more interestingly, the oscillations disappear when the V_{bias} increases, what is also not consistent with the observations in Figure 3.11. After thorough tests of the pixel, it turned out that the photodiodes where not sufficiently reverse biased. From the conceptual explanation performed in Figure 2.9 we know that for low photocurrents the voltage on the cathode of the photodiode is smallest. Similarly for the complementary P-type photoreceptor the anode on the anode of the photoreceptor higher. As a result voltages V_{p1} and V_{p3} were to close to the voltage V_{p2}, (from Figure 3.20) resulting in moderate forward biased diodes. The currents injected by the forward biased diodes caused oscillations in the photoreceptors. Adjustments of V_{bias} and V_{bias_ptype} helped to ensure that the diodes are properly reverse biased what solved the problem of the instability.

Chapter 5

Conclusion

The main goal of this thesis was to build and test the first focal plain array of bio-inspired AER pixels which mimic color opponencies between three primary colors in the retina.
The vision sensor was implemented in the TSMC 90nm CMOS process and tested by different measurement setups. The stacked photodiodes were characterised in terms of spectral response and proved to be capable of sensing three different color spectra. Although the test photodiodes provided much more photocurrent for long wavelengths rather than the short wavelengths, the photocurrent ratios could be used to extract information about color contrast changes. The implemented logarithmic photoreceptors and differencing amplifiers differentiated and amplified those photocurrent ratios resulting in streams of four different types of events/spikes. We managed to find unique spiking behaviour for different color transitions and we assigned color change meanings to those events. As a result the pixel spikes were translated directly into information, whether the incident light is *bluer*, *greener* or *redder*. Such behaviour is similar to the neuronal rate coding communication method: for the increase of the intensity of a stimulus, the frequency of a neuron *spike firing* increases. A simple visualizing algorithm plotting images in color was implemented and different color rotating objects were tracked and plotted as 2D images with a time scale. The speed of the *AER* peripheries used in this work was limited to $10 Meps$ for pixels of different rows, what was more than enough for the implemented array. During normal operation number of events rarely exceeded 2000 events per second. We managed to provoke a saturation of the handshake circuitry by stimulating all pixels at once and observe the *greediness* of the arbiters. However for normal operation conditions all pixels are not stimulated simultaneously and the saturation does not occur. Due to abandoning the frame-based read-out architecture, the vision sensor could track fast moving objects and its speed was only limited by a single pixel bandwidth which is much higher than the frame rate in traditional cameras. Also, the post processing load on a CPU performing object recognition based on colour code or coloured object outlines is much relaxed, because each pixel communicates individually and asynchronously only if it detects temporal colour contrast. Otherwise the pixels remain *silent*.
The results have been submitted for publication in form of a conference paper and a linked demonstration proposal (both included in Appendix E and F):

- Ł. Farian, J. A. Leñero-Bardallo, and P. Häfliger, *A bio-inspired AER temporal tri-color differentiator*, Submitted to IEEE Biomedical Circuits and Systems Conference (BioCAS), Lausanne, Switzerland. October 2014.
- Ł. Farian, J. A. Leñero-Bardallo, and P. Häfliger, *Live Demonstration: A bio-inspired AER temporal tri-color differentiator*, Submitted to IEEE Biomedical Circuits and Systems Conference (BioCAS), Lausanne, Switzerland. October 2014.

The tests of the chip revealed few issues:

- The analog part located close to the photodiodes was not properly shielded from the incident light. As a result, when the chip was stimulated by high intensity static blue light, the pixel generated unexpected spikes proportional to the light intensity. The blue light has highest dispersion coefficient, meaning that it reflects and dissipates most on the chip surface from all the wavelengths of the visible spectrum. Probably, blue light propagated and recombined near to the channel of the transistor of the differencing amplifier and resulted in injected photocurrent which provoked *fake* spikes. Another explanation of such behaviour may be an issue with wrong layout and a parasitic capacitance at the input of the circuit. This problem applied only to the single test pixel, not pixels in the array. The test pixel layout differs from the layout of the pixels implemented in the array because the test pixel has probes connected to its sensitive nodes. This could lead to increase of parasitic capacitances and make the source follower to charge the input of the differencing amplifier and increase its output.
- We observed pixels to be unstable for low light conditions. This problem was solved by increasing reverse voltage on the photodiodes, as described in the previous section.
- Some digital buffers did not send the pixels addresses or they had very limited output swing. As a result, it was impossible to read out the pixels addresses from the *AER* peripheries and display an image. The problem turned out to be not sufficient power distribution for analog buffers in the test pixel. They consumed to much power, resulting in substantial voltage drop on the power net what affected the digital buffers using the same net. The problem was solved by turning off the analog buffers.

This vision sensor has good potential for automotive robotics applications where high speed temporal color contrast recognition, especially of fast moving objects, is critical. We observed very strong response of the camera to blinking light sources of different colors. Such behaviour opens up possibilities to the interesting application for the camera, for instance to recognise the traffic lights colours in real time. Traffic lights are realised nowadays by high frequency blinking LEDs, which may be easily recognised by this camera. High dynamic range of this camera makes it even more suitable for such applications. For the frame-based cameras, on the other hand, it is possible that the frame rate is in phase with the frequency

of the stimuli and the stimuli may not be recorded at all. The power consumption of the implemented vision sensor is much smaller than the typical power consumption of traditional vision sensors. This is another important advantage of the sensor in terms of automotive applications.

5.1 Further Work

The future work considerations:

- The implemented visualizing algorithm translating events into color changes is very simple and it does not take into consideration different sensitivities of the stacked photodiodes nor the substrate coupling between pixels. Hence, future work shall involve more research to properly represent color changes in a reconstructed video stream by improving the visualizing algorithm.
- Since the light sensing in TSMC 90nm CMOS process was not explored, we decided to implement reasonably big photodiodes for test purposes. The photodiodes size was approximately $45x45\mu m$. After a spectral characterisation of these photodiodes, we know that also smaller photodiodes size would provide still reasonable photocurrents. Smaller size could help to improve the future camera resolution.
- More research may be required to study how to limit the substrate coupling which leads to the current injection into the PN junction and the *cross-talk* in the array.
- The charge injection from the reset transistor may be limited by adding a dummy complementary switch in parallel, which would help to prevent this charge from *traveling* to the sensitive input of the amplifier.
- The bandwidth of the reset transistor can be improved by limiting the swing of the gate voltage used to turn on and turn off this transistor.

Bibliography

[1] S. Haykin, *Neural Networks: A Comprehensive Foundation*. NJ, USA: Prentice Hall PTR: Upper Saddle River, second edition ed., 1998.

[2] F. H. Netter, *Atlas of Human Anatomy*. Arsley, USA: Ciba-Geigy Corp., 1989.

[3] B. A. Wandell, *Foundations of vision / Brian A. Wandell*. Sunderland, Mass. : Sinauer Associates, 1995. Includes bibliographical references (p. 445-462) and indexes.

[4] J. Moran and R. Desimone, "Selective attention gates visual processing in the extrastriate cortex," *Science, New Series*, 1985.

[5] M. A. Mahowald, "VLSI analogs of neuronal visual processing: a synthesis of form and function," *Computation and Neural Systems, Caltech, Pasadena, California*, pp. 27–44, 1992.

[6] C. A. Mead and M. Mahowald, "A silicon model of early visual processing," *Neural Networks*, vol. 1, no. 1, pp. 91 – 97, 1988.

[7] K. Boahen and P. Bioengineering, "A retinomorphic chip with parallel pathways: Encoding," in *INCREASING, ON, DECREASING, and OFF visual signals, Analog Integrated Circuits and Signal Processing*, pp. 121–135, Press, 2002.

[8] J. Lazzaro, J. Wawrzynek, M. Mahowald, M. Sivilotti, and D. Gillespie, "Silicon auditory processors as computer peripherals," *Neural Networks, IEEE Transactions on*, vol. 4, pp. 523–528, May 1993.

[9] E. Grenet, S. Gyger, P. Heim, F. Heitger, F. Kaess, P. Nussbaum, and P.-F. Ruedi, "High dynamic range vision sensor for automotive applications," *Proc. SPIE*, vol. 5663, pp. 246–253, 2005.

[10] P. Lichtsteiner, C. Posch, and T. Delbruck, "A 128×128 120dB $15\mu s$ latency asynchronous temporal contrast vision sensor," *IEEE Journal of Solid-State Circuits*, vol. 43, pp. 566–576, February 2008.

[11] M. Mahowald, *An Analog VLSI System for Stereoscopic Vision*. Norwell, MA, USA: Kluwer Academic Publishers, 1994.

[12] K. Boahen, "The retinomorphic approach: Pixel-parallel adaptive amplification, filtering, and quantization," *Analog Integrated Circuits and Signal Processing*, vol. 13, no. 1-2, pp. 53–68, 1997.

[13] J. A. Leñero-Bardallo, T.Serrano-Gotarredona, and B. Linares-Barranco, "A five-decade dynamic-range ambient-light-independent calibrated signed-spatial-contrast aer retina with 0.1-ms latency and optional time-to-first-spike mode," *Circuits and Systems I: Regular Papers, IEEE Transactions on*, vol. 57, pp. 2632–2643, Oct 2010.

[14] J. A. Leñero-Bardallo and P. Häfliger, "A dual operation mode bio-inspired vision sensor," *Biomedical Circuits and Systems Conference (BioCAS), 2013 IEEE*, pp. 310–313, Oct 2013.

[15] H. Okawa and A. P. Sampath, "Optimization of single-photon response transmission at the rod-to-rod bipolar synapse," *Physiology*, vol. 22, no. 4, pp. 279–286, 2007.

[16] J. Tiffin and H. S. Kuhn, "Color discrimination in industry," *Archives of Ophthalmology*, vol. 28, pp. 851–859, November 1942.

[17] J. M. A. Olsson and P. Hafliger, "Two color asynchronous event phot pixel," in *ISCAS*, pp. 169–172, 2008.

[18] J. A. Leñero-Bardallo, D. H. Bryn, and P. Häfliger, "Live demonstration: A bio-inspired asynchronous pixel event tri-color vision sensor," in *Circuits and Systems (ISCAS), 2012 IEEE International Symposium on*, pp. 726–726, May 2012.

[19] J. A. Leñero-Bardallo, D. H. Bryn, and P. Häfliger, "Flame monitoring with an aer color vision sensor," in *IEEE International Symposium on Circuits and Systems, ISCAS 2013, Beijing, China*, 2013.

[20] J. E. Dowling, *The Retina: An Approachable Part of the Brain*. The Belknap Press of Harvard University Press, 2 ed., 2012.

[21] R.Berner, P. Lichtsteiner, and T. Delbruck, "Event-based color change pixel in standard CMOS," in *ISCAS*, 2010.

[22] R. Berner and T. Delbruck, "Event-based pixel sensitive to changes of color and brightness," *IEEE Transactions On Circuits and Systems I*, vol. 58, no. 7, pp. 1581–1590, 2011.

[23] A. Darmont and Aphesa, "Spectral response of silicon image sensors," *Analog Integrated Circuits and Signal Processing*, 2009.

[24] J. K. e. a. Shih-Chii Liu, *Analog VLSI: Circuits and Principles*. Cambridge, Massachusetts, London, England: A Bradford Book, The MIT Press, 2002.

[25] R. F. Lyon and P. M. Hubel, "Eyeing the camera: Into the next century," in *10th Color Imaging Conf.*, pp. 349–355, 2002.

[26] "What's a pixel." `http://www.foveon.com/article.php?a=68`. Accessed: 2014-07-01.

[27] E. A. Vittoz, "Analog VLSI signal processing: Why, where, and how?," *Journal of VLSI signal processing systems for signal, image and video technology*, vol. 8, no. 1, pp. 27–44, 1994.

[28] C. C. Enz, F. Krummenacher, and E. A. Vittoz, "An analytical mos transistor model valid in all regions of operation and dedicated to low-voltage and low-current applications," *Analog Integr. Circuits Signal Process.*, vol. 8, pp. 83–114, July 1995.

[29] M. C. Delbrück, T., "Analog VLSI phototransduction by continuous-time, adaptive, logarithmic photoreceptor circuits," tech. rep., California Institute of Technology, Pasadena, CA, 1994.

[30] B. R., "Building-blocks for event-based vision sensor," tech. rep., ETH Zurich, 2011.

[31] S. Herculano-Houzel, "The human brain in numbers: a linearly scaled-up primate brain," *Frontiers in Human Neuroscience*, vol. 3, no. 31, 2009.

[32] P. H. Juan A. Leñero Bardallo, *Neuromorphic Electronics: Lecture Notes*, Fall 2012, [Available at]: `http://www.uio.no/studier/emner/matnat/ifi/INF5470/h12/timeplan/index.html`.

[33] D. Purves, G. J. Augusstine, and et al., *Neuroscience*. Sunderlan: Sinauer Associates, Inc., fourth edition ed., 2008.

[34] P. Häfliger, *A spike based learning rule and its implementation in analog hardware*. PhD thesis, ETH Zürich,Switzerland, Sunderlan, 2000.

[35] J. Leñero Bardallo, D. Bryn, and P. Häfliger, "Bio-inspired asynchronous pixel event tricolor vision sensor," *Biomedical Circuits and Systems, IEEE Transactions on*, vol. PP, no. 99, pp. 1–1, 2013.

[36] B. Razavi, *Design of Analog CMOS Integrated Circuits*. New York, NY, USA: McGraw-Hill, Inc., 1 ed., 2001.

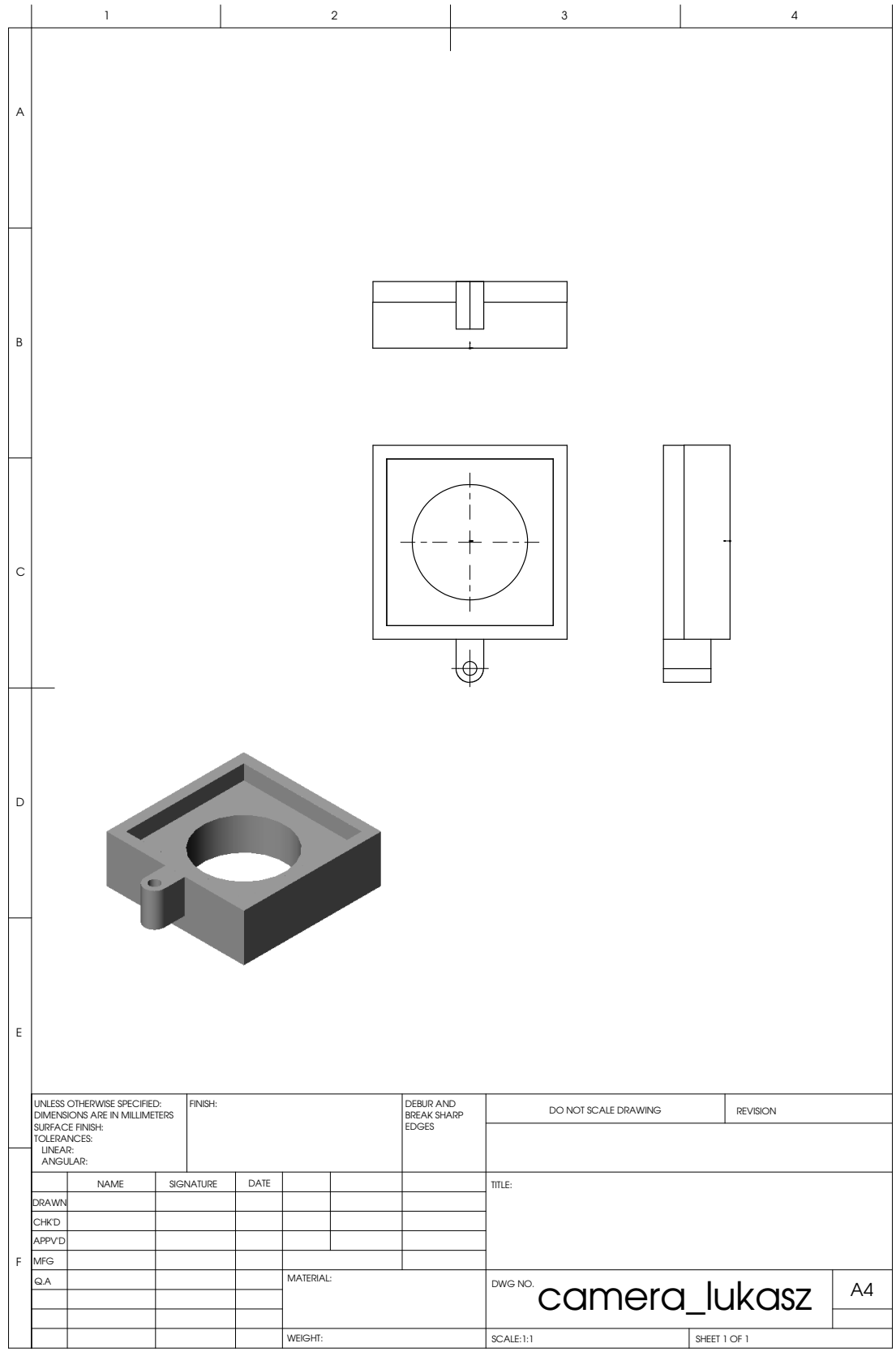
[37] T. C. Carusone, K. Martin, and D. Johns, *Analog integrated circuit design*. Hoboken, N.J.: Wiley, 2013. 2nd ed., International student ed.

[38] R. Berner, T. Delbruck, A. Civit-Balcells, and A. Linares-Barranco, "A 5 meps $100 usb2.0 address-event monitor-sequencer interface," in *Circuits and Systems, 2007. ISCAS 2007. IEEE International Symposium on*, pp. 2451–2454, May 2007.

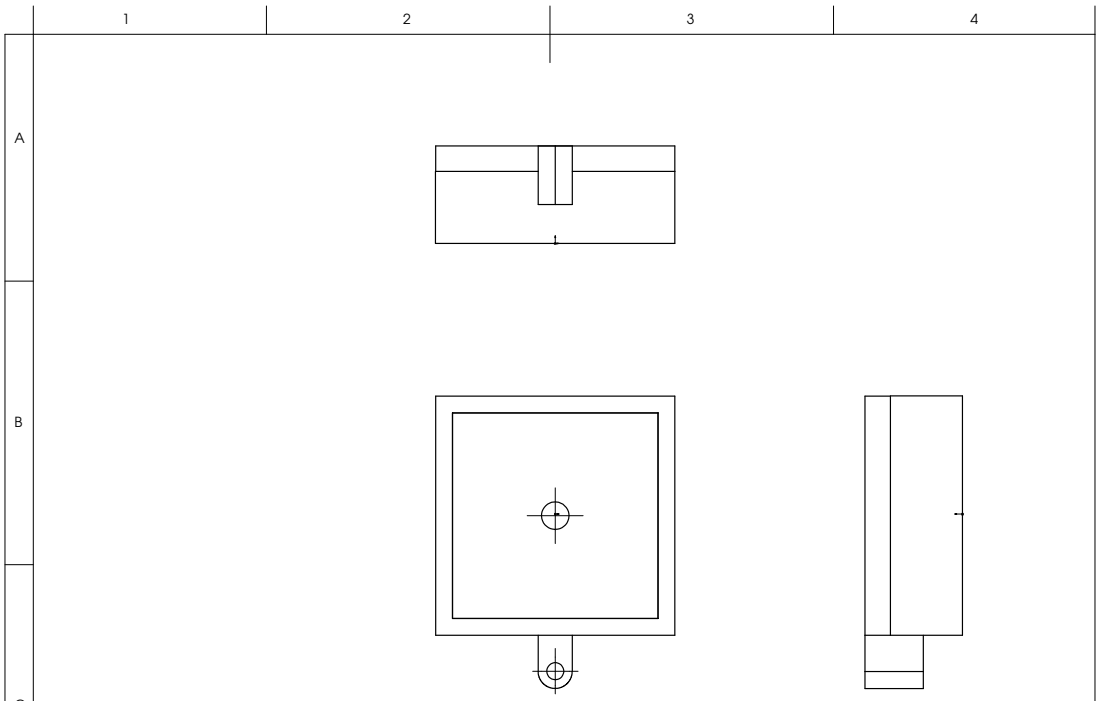
[39] jAER Open Source Project. [online], "Available: `http://www.sourceforge.net/apps/jaer/wiki`."

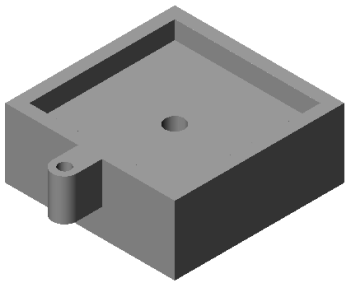
Appendix A

Drawings in SolidWorks

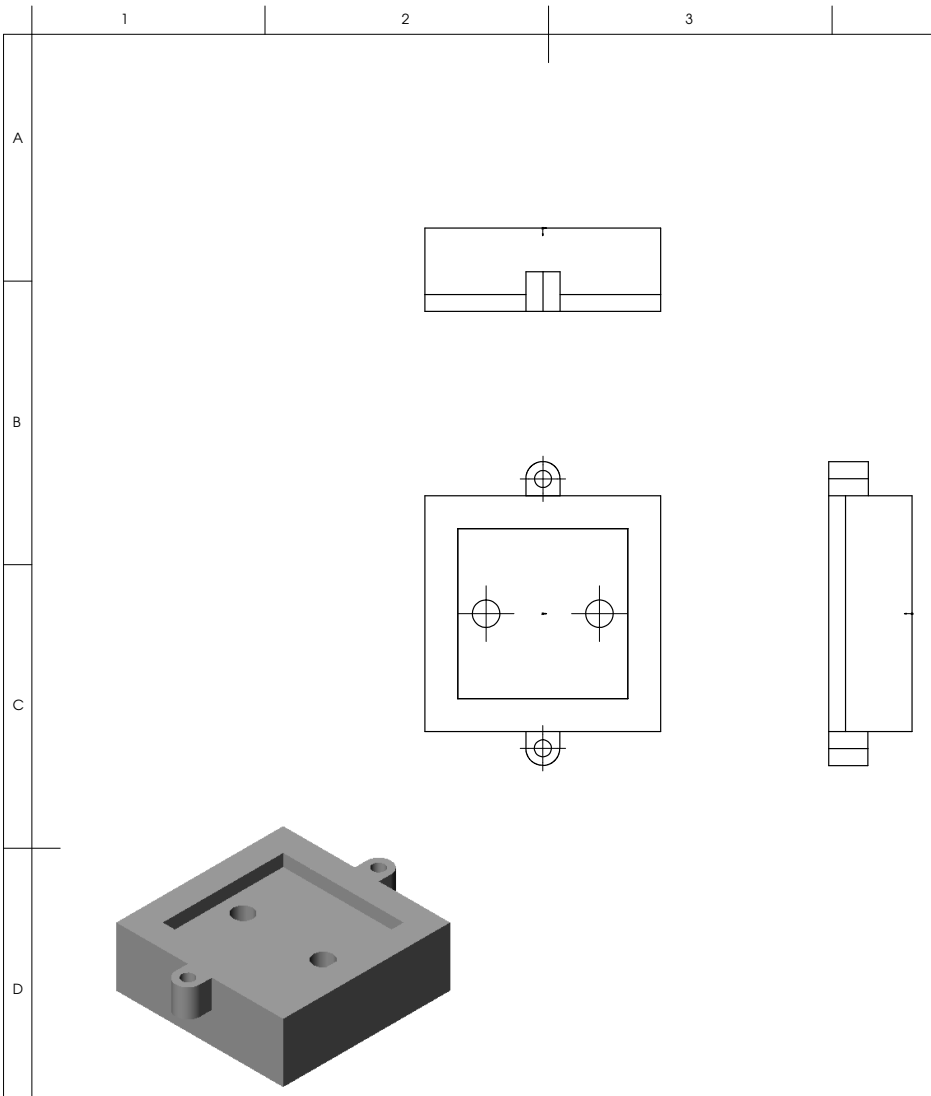


UNLESS OTHERWISE SPECIFIED: DIMENSIONS ARE IN MILLIMETERS SURFACE FINISH: TOLERANCES: LINEAR: ANGULAR:	FINISH:		DEBUR AND BREAK SHARP EDGES	DO NOT SCALE DRAWING	REVISION	
	NAME	SIGNATURE	DATE			TITLE:
DRAWN						
CHK'D						
APPV'D						
MFG						
Q.A				MATERIAL:	DWG NO. camera_lukasz	A4
				WEIGHT:	SCALE:1:1	SHEET 1 OF 1





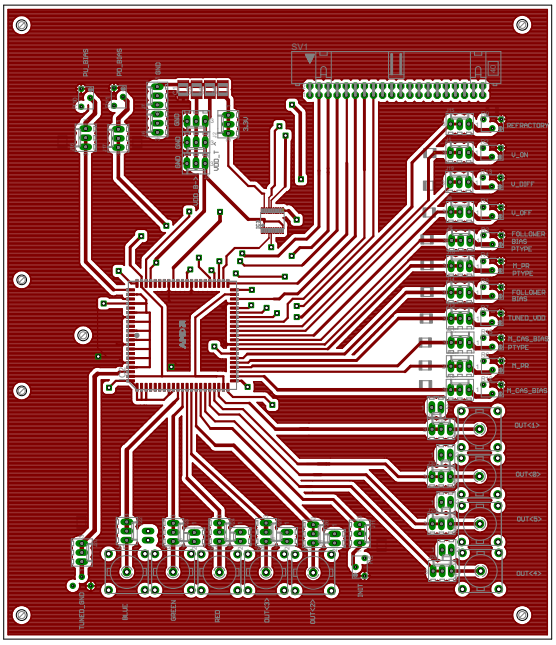
UNLESS OTHERWISE SPECIFIED: DIMENSIONS ARE IN MILLIMETERS SURFACE FINISH: TOLERANCES: LINEAR: ANGULAR:		FINISH:			DEBUR AND BREAK SHARP EDGES	DO NOT SCALE DRAWING	REVISION
	NAME	SIGNATURE	DATE			TITLE:	
DRAWN							
CHK'D							
APPV'D							
MFG							
Q.A				MATERIAL:		DWG NO. camera_lukasz_fibre	A4
				WEIGHT:		SCALE:1:1	SHEET 1 OF 1

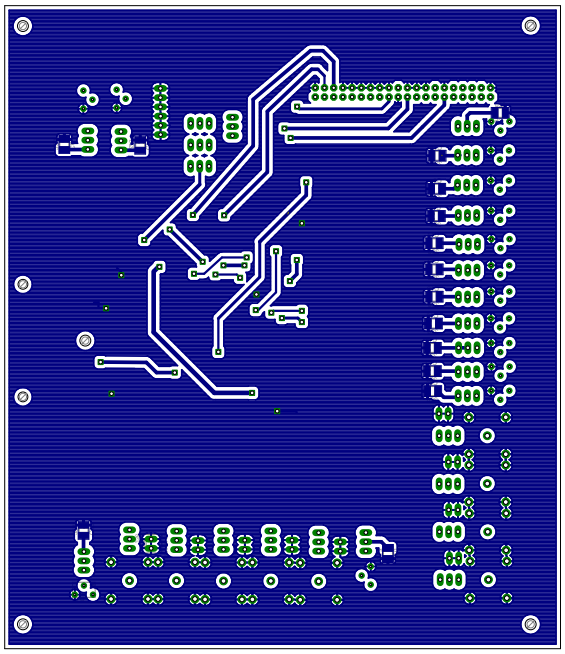


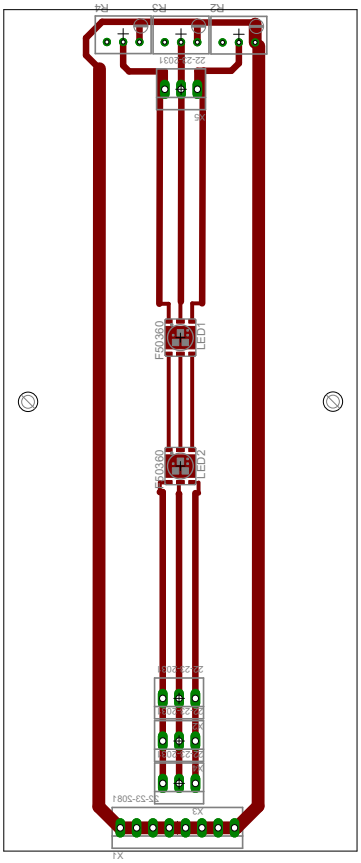
UNLESS OTHERWISE SPECIFIED: DIMENSIONS ARE IN MILLIMETERS SURFACE FINISH: TOLERANCES: LINEAR: ANGULAR:		FINISH:		DEBUR AND BREAK SHARP EDGES	DO NOT SCALE DRAWING	REVISION
	NAME	SIGNATURE	DATE		TITLE:	
DRAWN						
CHK'D						
APPV'D						
MFG						
Q.A				MATERIAL:	DWG NO: LED_double_lukasz	A4
				WEIGHT:	SCALE:1:1	SHEET 1 OF 1

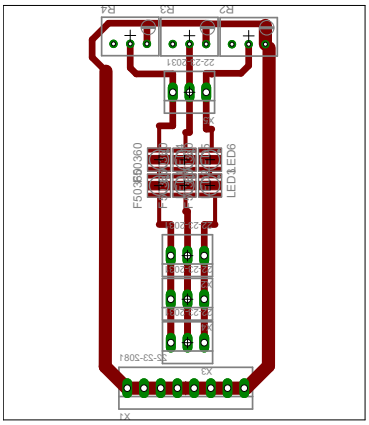
Appendix B

Testing PCBs layout









Appendix C

Layout and microscope photos of the chip

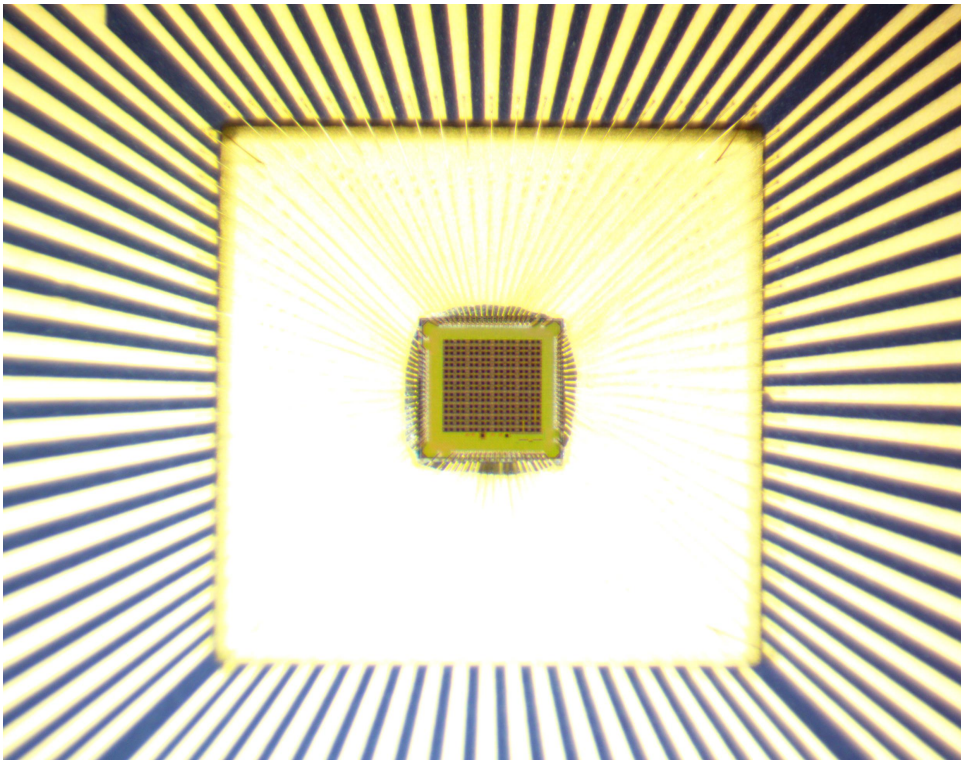
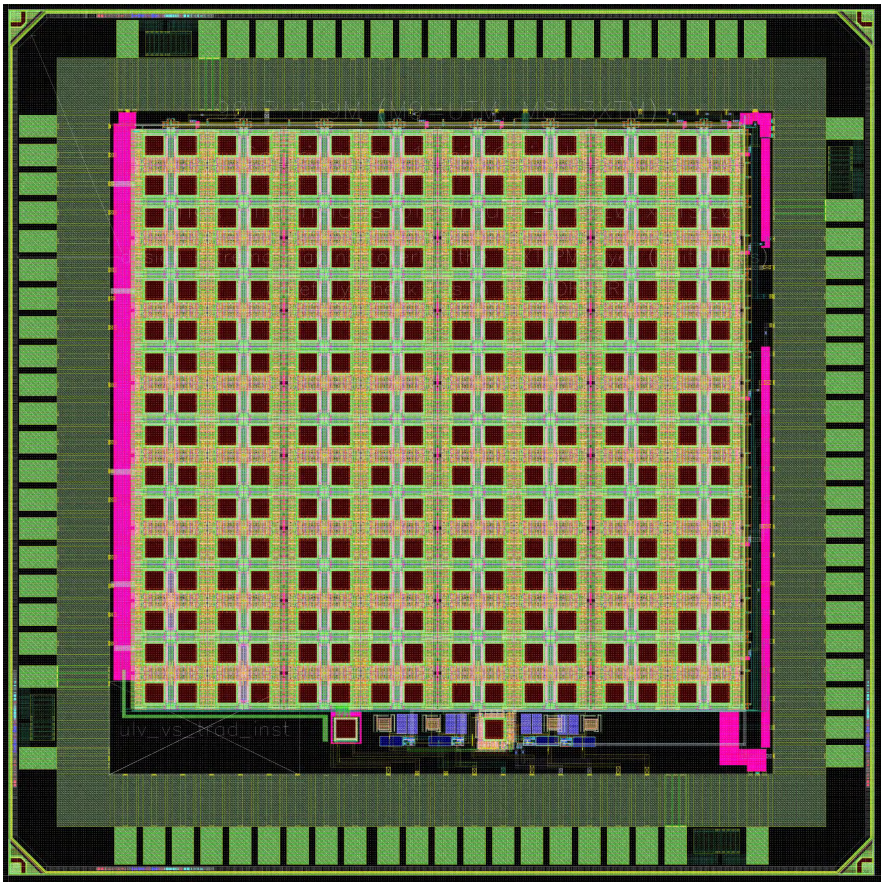
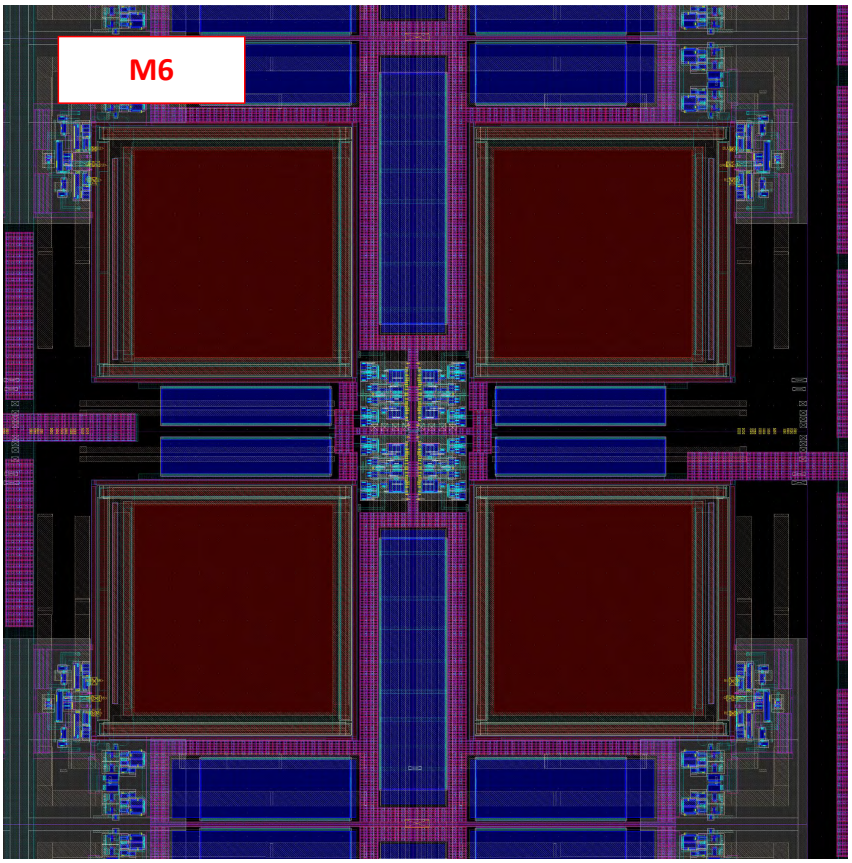
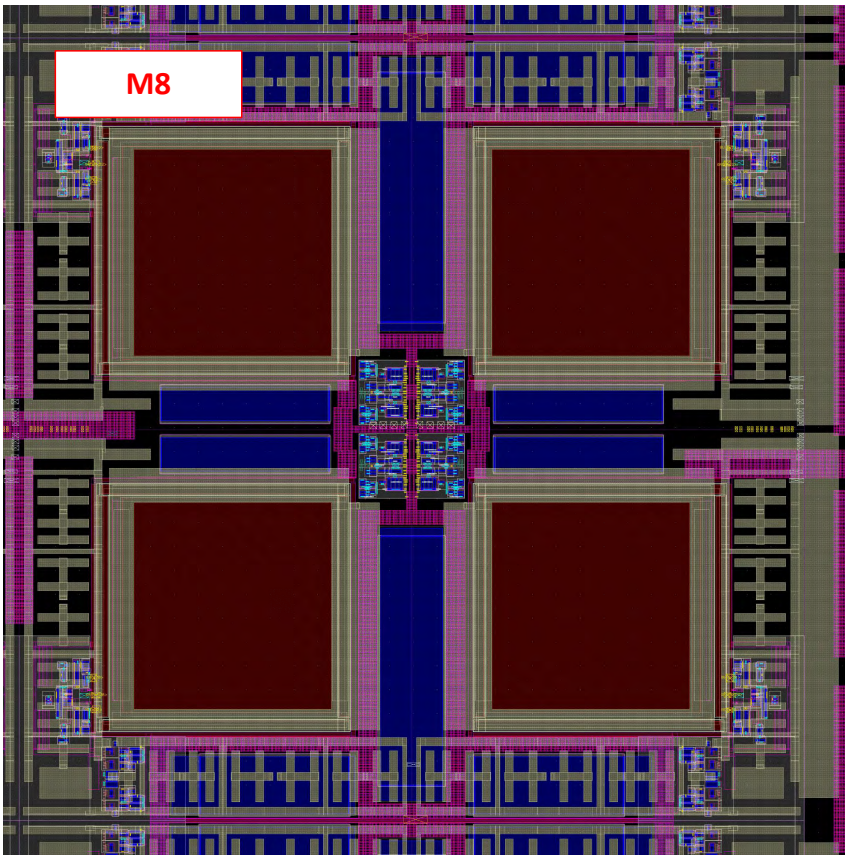


Fig. C.1: Vision sensor chip with bond wires threaded to the pads



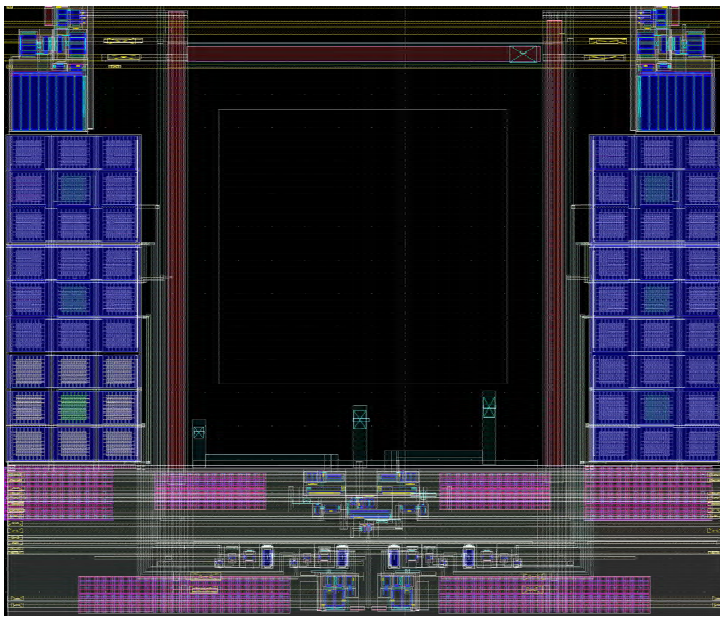

M6


M8

Appendix D

Alternative layout of the pixel

The alternative pixel layout including MOM capacitors.



Appendix E

BioCAS 2014 paper: A Bio-Inspired AER Temporal Tri-Color Differentiator

A Bio-Inspired AER Temporal Tri-Color Differentiator

Łukasz Farian, Juan Antonio Leñero-Bardallo, Philipp Häfliger
Department of Informatics, University of Oslo, Norway, email: lukaszf@ifi.uio.no

***Abstract*—The paper presents the first focal plane array of asynchronous event pixels that react to temporal color contrast of three different color spectra. The three different spectra are transduced into photo currents by stacked photo diodes and temporal changes of the contrast of these three spectra are quantified as pulse density modulated signals and conveyed off-chip by the Address Event Representation (AER) protocol. 16x16 pixel array has been implemented and tested in a 90nm TSMC CMOS process. A single pixel measures 82μm x 82μm and has a fill factor of 27% and a dynamic range of 5 decades. The paper presents single pixel characteristics and shows visualization examples of the pixel array.**

I. Introduction

Color contrast detection is important in many industrial processes where robots are involved. For instance, it is essential to match samples of various colored materials or to operate mobile equipment for which the operating signals are colored lights [1].

Conventional color image sensors are frame-based and use color filters to extract color information from the visual scene. Filters are placed over each pixel. Typically, they are band-pass filters centered a red, green, and blue colors (RGB), usually arranged in groups of four filters (RGBG) called Bayer masks.

In the last years, we have seen an increase of bio-inspired vision sensors that try to emulate some features of the human retina [2]–[5], just to mention but a few. Typically they detect spatio-temporal contrast or light intensity. However, such sensors have been traditionally limited to black-and-white sensors. This limitation is mainly due to their low resolution arrays. Placing discrete color filters over pixels implies reducing the sensor resolution at least three times. RGB color processing with bio-inspired vision sensors is recent [6] and authors prefer stacked photodiodes at different depths to sense the color information instead of discrete optic filters.

However, neural processing of color is not just based on filtering light at three different wavelengths. We do have cones with three different pigments that can filter light approximately at red, green, and blue colors, but we do not perceive colors as separate hues [7]. Certain pairs of colors are mutually exclusive. Actually there is an opponency between red and green and between blue and yellow colors. To our knowledge, the neural full color processing has not been modeled with vision sensors yet.

This work presents a two-dimensional pixel array whose pixels spike when they detect transitions between three different color spectra, trying to emulate the color opponencies in the human retina. The sensor is a prototype with stacked photodiodes [6] with different spectral sensitivity. Outputs provide information about the different color transitions. The sensor can detect transitions between colors at high speed, which generate asynchronous pixel events that are conveyed off-chip by the Address Event Representation (AER) protocol. Possible applications are for example automated industrial processes where transitions between colors need to be detected for appropriate processing [1], automotive systems, or even systems that need to detect transitions between radiation in the near-infrared band and the visible spectrum [8].

II. Operation Principle

The designed pixel monitors three different color spectra and it reports whenever the ratio of the light intensity filtered by those three spectra changes. For illustration you may think of those three spectra as red, green, and blue light and denote the filtered intensities R, G, and B. Each pixel reports quantitative changes in the ratios R/G and G/B. The third ratio R/B can be indirectly obtained from the relationship R/B=R/G*G/B. A change in spectrum is reported as an asynchronous pixel event on the AER bus: if R/G has increased the pixel produces positive R/G events on the AER bus. If G/B has decreased, the pixel reports a negative G/B events. The density of these events is proportional to the rate of the change. Please note that if the light level of the pixel changes without a change in the spectrum, the pixel will not react at all. Thus, its output will be independent from the illumination level of the input. Later we will refer to the changes in spectra (R, G and B) and photo currents (I_R, I_G and I_B) interchangeably because of the linear dependence between these two quantities [9].

III. Vision Sensor Design

This section describes the design of a single pixel and a focal plain array.

A. Pixel Design

Fig. 1 shows a block diagram of a single pixel. It can be divided into three stages: photoreceptor, differencing amplifier and comparators. The pixel topology can be treated as an extension of the design implemented by Berner and Delbrück in [10]. Their pixel comprises two stacked photo diodes (buried double junction) providing two different color spectra, whereas we implemented a pixel made up by 3 different photo diodes stacked at different depths, providing three different color spectra [11]. The cross section of the implemented stacked photo diodes in Fig. 2 shows that these photo diodes

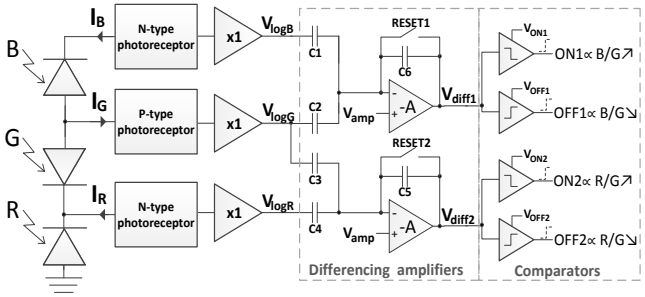


Fig. 1: Block diagram of a single pixel.

are connected in series. We can assume that they provide three photo currents: I_R, I_G and I_B proportional to the affore named R, G and B spectra respectively [9]. These photocurrents are the result of the different absorption spectra of the stacked photo diodes which depend on their depths in the substrate. The 'B' spectrum is the spectrum of the top diode, the 'G' spectrum is the sum of the spectra of the middle and top diode and the 'R' spectrum is the sum of the middle and bottom diode.

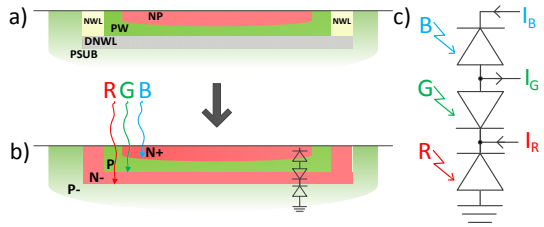


Fig. 2: a) Cross section of the stacked photodiodes, b) Formation of PN junctions, c) Equivalent schematic of the stacked photodiodes.

The photocurrents are converted into voltages V_{logx} by a well known logarithmic photoreceptor [12], having high dynamic range and extended bandwidth. The middle p-type photoreceptor *sinks* the input photo current [10] and hence it has opposite gain compared to the other two photoreceptors. Since the differencing amplifiers remove the DC level, the

Event	Spectra ratio	Color change
ON1	$\propto \frac{B}{G}$ ↗	bluer
OFF1	$\propto \frac{B}{G}$ ↘	greener
ON2	$\propto \frac{R}{G}$ ↗	redder
OFF2	$\propto \frac{R}{G}$ ↘	greener

TABLE I: Translation of ONx, OFFx events into color changes.

resulting output voltages ΔV_{diff1} and ΔV_{diff2} will be the amplified differences $\Delta(V_{log1} - V_{log2})$ and $\Delta(V_{log3} - V_{log2})$ respectively. Assuming the gains in all branches are the same [10], ΔV_{diff1} and ΔV_{diff2} become respectively:

$$\Delta V_{diff1} = -A_1 \cdot U_T \cdot \Delta ln\left(\frac{I_B}{I_G}\right) \propto \Delta\frac{B}{G} \quad (1)$$

$$\Delta V_{diff2} = -A_2 \cdot U_T \cdot \Delta ln\left(\frac{I_R}{I_G}\right) \propto \Delta\frac{R}{G}, \quad (2)$$

where $A_1 = C_{1/2}/C_6$ and $A_2 = C_{3/4}/C_5$ are the gains of the differencing amplifiers determined by the ratio between the input and the feedback capacitors. Equations 1 and 2 indicate that the pixel responds only to changes in the spectra ratios: R/G and B/G. If the light level of the pixel changes without a change in the spectrum, the pixel will not react because the ratio remains constant. Thus, the pixel is, ideally, independent from the illumination level of the input. Please note that the differencing amplifier is actually inverting. As a result, if for instance B/G ratio increases, then V_{diff1} drops, and if V_{diff1} exceeds the adjustable threshold V_{ON1}, the pixel produces positive event ON1. Conversely, if the ratio decreases, the pixel produces OFF1 event. Similarly changes in the second ratio R/G are indicated by ON2 and OFF2 events. After event is generated the differencing amplifier is reset by RESETx signal and the process repeats. Table I explains how the events are translated into color changes.

B. Matrix and arbitration

Since each single pixel can generate 4 different events, and we implemented 256 pixels, there are 1024 different addresses to be handled by the arbitration system. We chose the AER arbitration system proposed by Häfliger in his PhD [13]. It is capable of handling event rates up to $10Meps$, which proved to be sufficient for our purposes.

IV. Experimental Results

We implemented the 16x16 pixels array in the TSMC 90nm CMOS process. Fig.3 shows a designed custom PCB with a $4mmF/1.2$ C-mount lens, the layout of a single pixel and a micrograph of the chip. The AER data from the vision sensor is logged by the USBAERmini2 board [14] and then this data is sent through a USB port to a PC.

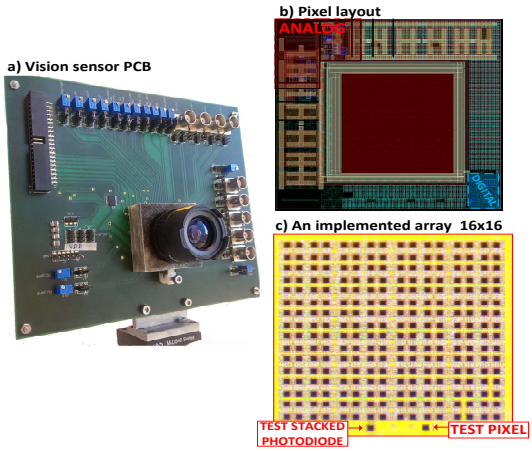
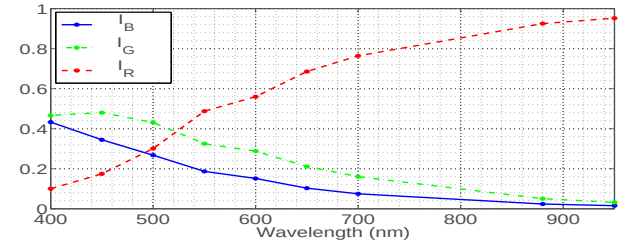


Fig. 3: a) Designed custom PCB, b) layout of a single pixel, size of the pixel 82μm x 82μm, fill factor: 27%, digital area occupies: ≈1%, analog part: 13%, c) Microscope photo of the implemented vision sensor.

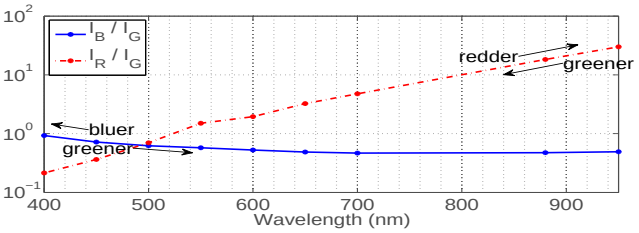
A. Stacked Photo diode

Fig. 4a shows the three measured normalized photo currents of one pixel when illuminated by the Optical Feedback Controlled Tungsten Halogen Light Source with color filters

centered between 400 nm and 950 nm. Normalized means here that the sum of the currents is normalized to 1.0. The true spectra of the pixel do indeed have center frequencies ranging from the lowest for what we called 'R' to the highest that we called 'B', but are in truth much more wide band than true RGB filters. So for example a narrow band light source moving from 500nm to 600nm wavelength would clearly lead to positive R/G events. More subtly, it would also lead to a few positive G/B events, as the green and the blue line decrease in parallel on this linear scale, that means their G/B *ratio* will increase. Fig.4b shows the calculated photo



(a) Photo currents relative values.



(b) Photo currents ratio,

Fig. 4: Spectral response of 3 stacked photo diodes: a) relative values of their photo currents, b) photo currents ratios amplified by the differencing amplifiers (Equations 1,2). The values were corrected by the Halogen Light Source irradiance factor and color filters transmission efficiency.

current ratios and illustrates how these ratios can be used to discriminate between colors: the R/G ratio increases when a the dominating light wavelength increases (the color spectrum becomes *redder*) and B/G decreases when the dominating light wavelength decreases (the color spectrum becomes *bluer*). The ratio B/G exhibits weaker dependence on color spectra, it has steepest slope up to 650 nm and after that is flat, while the R/G ratio has almost constant slope over the entire tested spectrum. Thus if for instance a narrow band light source moves from 600 nm to 900 nm wavelength, it would lead to positive R/G events and a few negative B/G events.

B. Color recognition

We used high-power full color LEDs F50360 from Seoul Semiconductor to characterize a single pixel response to changes in the ratio of light intensities. These photodiodes emit light centered in 625nm, 525nm and 460nm wavelengths.

First, we characterized the pixel response to a continuous red-green intensity change by illuminating its surface with 70Hz sinusoidal out-of-phase red and green lights. Fig.5 shows the current through the LEDs and the response of a single pixel to the stimulus. We observed two ON2 and two OFF2 events per cycle. Since we were changing the R and G spectra inversely, the change of the R/G ratio is most substantial and hence the output ΔV_{diff2} exceeds the V_{ON2} and V_{OFF2} thresholds. The V_{ONx} threshold was $\approx 1.2V$ and V_{OFF} threshold $\approx 1.3V$

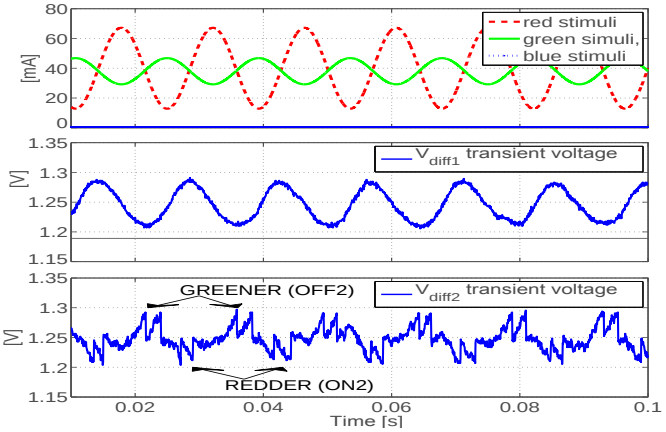


Fig. 5: Current consumption of the photodiodes and a single pixel response to the continuous *red-green* contrast change.

ΔV_{diff1} moves between threshold values, but does not exceed them and, hence, no ON1 nor OFF1 events are generated. Since there has not been any blue light change in the stimulus (the blue LED was off), the nominator and denominator in the B/G ratio change by the same amount, and ideally this ratio should remain constant. In fact, the output ΔV_{diff1} fluctuates only because the photo currents ratio B/G exhibits weak dependence on longer spectrum wavelengths, as was shown in Fig.4b. Next, we tested a response of the

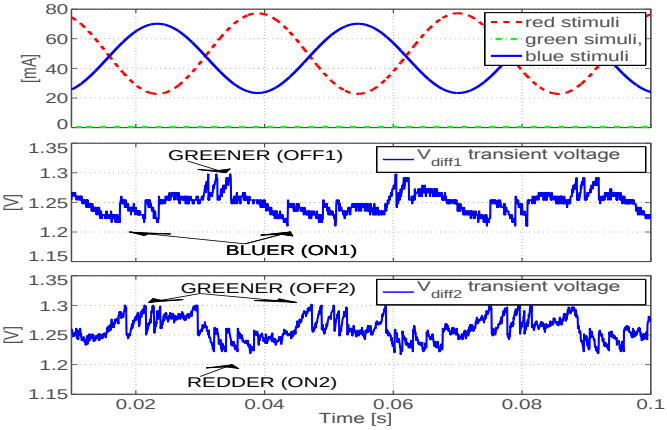


Fig. 6: Current consumption of the photodiodes and a single pixel response to the continuous *red-blue* contrast change.

pixel to the continuous *blue-red* contrast change (Fig.6). This time we observed 4 types of events. Events ON1 occur when the blue light positive transient is biggest (B/G increases) and conversely, ON2 events occur at the maximum red positive transient (R/G increases). Besides, we observed OFF1 and

OFF2 events that might be confusing at first glance. However, OFF1 events indicate a shift to longer wavelengths from blue and OFF2 events indicate a shift to shorter wavelengths from red and thus, the responses are as expected and correct.

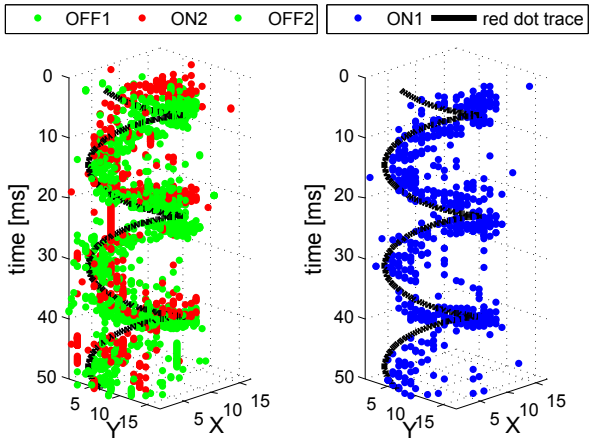


Fig. 7: Space-time plot of the rotating red dot on the green background. Number of events: ON1=924, OFF1=1107, ON2=5060, OFF2=5210

Fig.7 shows the response to a rotating red dot on a green background and Fig.8 shows the response to a rotating blue dot on a red background. According to Table I we assigned green color to OFF1 and OFF2 events, red color to ON2 events, and blue color to ON1 events. Comparing Fig.7 and Fig.8 we can observe that for a rotating red dot roughly 5 times more events representing change of R/G ratio are registered. For a rotating blue dot the vision sensor generates much less events representing change of R/G ratio. In Fig.7 the red events preceed the center of the red dot stimulus while the green and blue events are trailing. On the other hand, in Fig.8 the blue and green events preceed the center of the red dot, while red events are trailing. We repeated this experiment for different background light conditions, ranging from 0.1lux up to 15klux and we obtained similar behavior, meaning that the vision sensor has a dynamic range of > 5 decades.

V. Conclusion

We built and tested the first focal plain array of bio-inspired AER pixels which mimics full color opponencies between three primary colors in the retina. Its pixels are capable of detecting temporal contrast of three different color spectra sensed from three stacked photo diodes. We obtained experimental results from a single test pixel and from the array. Since the vision sensor abandons a global integration time and each pixel communicates individually and asynchronously whenever it detects temporal color contrast, the post processing load on a CPU performing object recognition based on color codes or colored object outlines is much relaxed. We implemented a simple visualizing algorithm translating events into color changes. Future work shall involve more research to properly represent color changes in a reconstructed video stream. This vision sensor has good potential for an automotive robotics applications where a high speed temporal color contrast recognition, especially of fast moving objects, is critical.

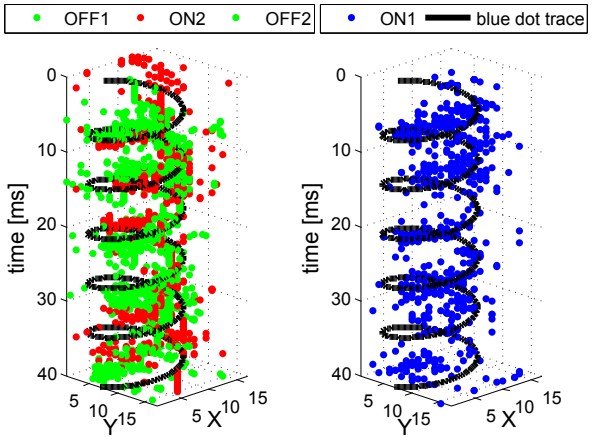


Fig. 8: Space-time plot of the rotating blue dot on the red background. Number of events: ON1=1208, OFF1=1353, ON2=2845, OFF2=2795

References

[1] J. Tiffin and H. S. Kuhn, "Color discrimination in industry," *Archives of Ophthalmology*, vol. 28, no. 5, pp. 851–859, November 1942.

[2] J. A. Leñero-Bardallo, T. Serrano-Gotarredona, and B. Linares-Barranco, "A 5-decade dynamic range ambient-light-independient calibrated signed-spatial-contrast aer retina with 0.1ms latency and optional time-to-first-spike mode," *Transactions on Circuits and Systems, Part-I*, 2010.

[3] C. Posch, D. Matolin, and R. Wohlgenannt, "A QVGA 143dB dynamic range asynchronous address-event PWM dynamic image sensor with lossless pixel-level video compression," *IEEE Journal of Solid State Circuits*, vol. 46, no. 1, pp. 259–275, January 2010.

[4] T. Serrano-Gotarredona and B. Linares-Barranco, "A 128 $\times$ 128 1.5% contrast sensitivity 0.9% FPN 3μs latency 4mW asynchronous frame-free dynamic vision sensor using transimpedance preamplifiers," *IEEE Journal of Solid-State Circuits*, vol. 48, no. 3, pp. 827–838, March 2013.

[5] P. Lichtsteiner, C. Posch, and T. Delbruck, "A 128 $\times$ 128 120dB 15μs latency asynchronous temporal contrast vision sensor," *IEEE Journal of Solid-State Circuits*, vol. 43, no. 2, pp. 566–576, February 2008.

[6] J. A. Leñero-Bardallo, D. H. Bryn, and P. Häfliger, "Bio-inspired asynchronous pixel event tri-color vision sensor," *IEEE Transactions on Biomedical Circuit and Systems*, pp. 1–13, August 2013.

[7] J. E. Dowling, *The Retina: An Approachable Part of the Brain*, 2nd ed. The Belknap Press of Harvard University Press, 2012.

[8] J. A. Leñero-Bardallo, D. H. Bryn, and P. Häfliger, "Flame monitoring with an aer color vision sensor," in *IEEE International Symposium on Circuits and Systems, ISCAS 2013, Beijing, China*, 2013.

[9] B. Tabbert and A. Goushcha, "Linearity of the photocurrent response with light intensity for silicon pin photodiode array," in *Nuclear Science Symposium Conference Record, 2006. IEEE*, vol. 2, Oct 2006, pp. 1060–1063.

[10] R. Berner and T. Delbruck, "Event-based pixel sensitive to changes of color and brightness," *IEEE Transactions On Circuits and Systems I*, vol. 58, no. 7, pp. 1581–1590, 2011.

[11] J. A. Leñero-Bardallo, D. H. Bryn, and P. Häfliger, "Bio-inspired asynchronous pixel event tri-color vision sensor," in *IEEE Transactions on Biomedical Circuit and Systems*, August 2013, pp. 1–13.

[12] M. C. Delbrück, T., "Adaptive photoreceptor with wide dynamic range," in *Circuits and Systems, 1994. ISCAS '94., 1994 IEEE International Symposium on*, vol. 4, May 1994, pp. 339–342 vol.4.

[13] P. Häfliger, "A spike based learning rule and its implementation in analog hardware," Ph.D. dissertation, ETH Zürich,Switzerland, Sunderlan, 2000.

[14] R. Berner, T. Delbruck, A. Civit-Balcells, and A. Linares-Barranco, "A 5Meps $100 USB 2.0 address-event monitor-sequencer interface," in *ISCAS 2007, New Orleans*, 2007, pp. 2451–2454.

Appendix F

BioCAS 2014 paper: Live Demo: A Bio-Inspired AER Temporal Tri-Color Differentiator

Live Demonstration: A Bio-Inspired AER Temporal Tri-Color Differentiator

Łukasz Farian, Juan A. Leñero-Bardallo, Philipp Häfliger
Nanoelectronics group, Department of Informatics, University of Oslo, Norway
E-mails: lukaszf@ifi.uio.no

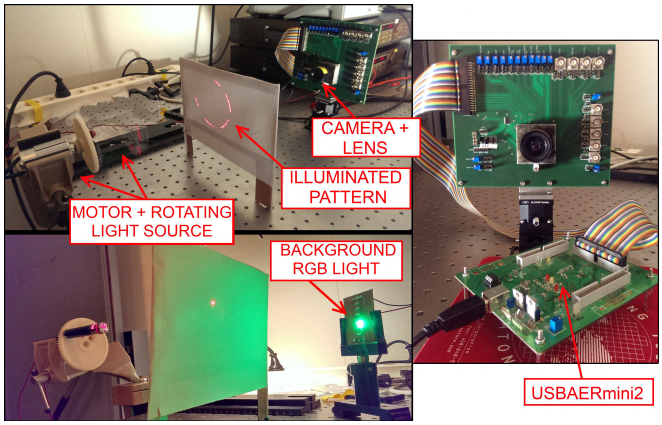


Fig. 1. Experimental setup.

***Abstract*—We demonstrate the first array of asynchronous event pixels that react to temporal color contrast of three different color spectra. The three different spectra are transduced into photo currents by stacked photo diodes. Temporal changes of the contrast of these three spectra are quantified as pulse density modulated signals and conveyed off-chip by the Address Event Representation (AER) protocol. The 16×16 pixel array has been fabricated in the standard TSMC 90nm CMOS process.**

I. Demonstration Setup

The demonstration system is composed of a laptop, 2 PCB boards, a light source attached to a motor and a white sheet of paper, as shown in Fig. 1. One PCB contains a lens mount and the sensor. The other one is a data logger USBAERmini2 [1] that stores the data and sends it via USB port to the laptop. The laptop has installed a dedicated Java interface [2] to display real-time color images.

II. Visitor Experience

The new sensor emulates the neural processing of color in the human retina. Neural processing of color is not just based on filtering light at three different wavelengths how commercial frame-based sensor do. Humans have cones with three different pigments that can filter light approximately at red, green, and blue colors, but we do not perceive colors as separate hues [3]. Certain pairs of colors are mutually exclusive. Actually there is an opponency between red and green and between blue and yellow colors.

This demonstration presents a two-dimensional pixel array whose pixels spike when they detect transitions between primary colors, trying to emulate the color opponencies in the human retina. The sensor is a prototype with stacked photo diodes with different spectral sensitivity [4]. Pixel output spikes provide information about the transient current variations of three stacked photo diodes with different spectral sensitivity.

The visitor will directly interact with the sensor and observe how temporal transitions between primary colors are detected. A moving light source will be projected on the white sheet of paper in front of the sensor. It will react depending on the incident light spectra. White light will not provoke output events. Colored light (blue, green or red) will make pixels spike out, because they detect transient spectra variations. Furthermore, a high frequency blinking light source (blue, green or red) will be placed in front of the sensor. The visitor will observe that the camera strongly responds to the high frequency color changes and generates many output events. Pixel outputs events will be displayed on a laptop screen using a custom Java interface. Furthermore, the visitors will be able to interact freely with the system and experiment how it reacts to the colored moving objects.

III. Reference Publications

The article of Farian et al. [5] describes the sensor in detail. It has been invited for presentation to the special session "Bio-Inspired Circuits and Architectures for Robotics" at BioCAS 2014. The board used to store and transmit the sensor information to a PC is described in [1]. A dedicated Java custom interface [2] was adapted and programmed to show real-time color transitions on a computer screen.

References

[1] R. Berner, T. Delbruck, A. Civit-Balcells, and A. Linares-Barranco, "A 5Meps $100 USB 2.0 address-event monitor-sequencer interface," in *ISCAS 2007, New Orleans*, 2007, pp. 2451–2454.
[2] "jAER open source project," http://sourceforge.net/projects/jaer/.
[3] J. E. Dowling, *The Retina: An Approachable Part of the Brain*, 2nd ed. The Belknap Press of Harvard University Press, 2012.
[4] J. A. Leñero-Bardallo, D. H. Bryn, and P. Häfliger, "Bio-inspired asynchronous pixel event tri-color vision sensor," *IEEE Transactions on Biomedical Circuit and Systems*, pp. 1–13, August 2013.
[5] L. Farian, J. A. Leñero-Bardallo, and P. Häfliger, "A bio-inspired AER temporal tri-color differentiator," *Submitted to IEEE Biomedical Circuits and Systems Conference (BioCAS), Lausanne, Switzerland. Preprint available at http://folk.uio.no/juanle/Downloads/BioCAS2014.pdf*, October 2014.